

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LA CONSCIENCE DE SOI ILLUSOIRE : DEREK PARFIT ET SES CRITIQUES

MÉMOIRE

PRÉSENTÉ COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN PHILOSOPHIE

PAR GABRIEL LACOSTE

AVRIL 2015

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je débiterai ce mémoire en remerciant tous les proches qui en ont accompagné la réflexion durant de longues années: Lysiane Goulet-Gervais, Mathieu Tremblay, Dominic Gagnon, Jean-Nicolas Paul, Julie Roussil, Vincent Perron, Maxime Lefebvre, Guillaume Fleurant, Hughes Jutras, Édith Plourde-Bérubé, Pierre-Yves Rochefort, Anton Uvarov, Marie-Lise Chrétien-Pineault, Gabriel Proulx, André-Anne Robert, Ève Salvas, Marise Daigle, Guy Rolland, Marc-André Langelier, Laura Darche, Caroline Guy, ma mère et ma sœur. C'est grâce à leur présence, ainsi qu'à leur réaction sympathique, hostile, sceptique, confuse ou indifférente que j'ai dû constamment ajuster mes hypothèses, mon langage et mes perceptions à une réalité extérieure plutôt que de me laisser errer librement vers des pistes qui ne mènent nulle part. Ce sont eux qui m'ont rendu profondément méfiant de la conscience que j'ai de moi-même. Je remercie également Denis Fisette, Mauro Rossi et Alain Voizard pour l'encadrement de mon projet. Ceux-ci m'ont permis de mesurer l'état de ma réflexion relativement à sa cohérence et à sa clarté. Je tiens également à mentionner au passage l'apport inestimable des usagers du programme l'Ancrage-Delta de l'Armée du Salut pour personnes aux prises avec des problèmes d'alcoolisme et de toxicomanie. J'y ai travaillé durant des années. Leurs expériences d'échec, d'adversité et de succès quant à la réalisation de l'image qu'ils ont d'eux-mêmes m'ont permis de comprendre les enjeux associés aux illusions de l'Égo avec une profondeur inestimable.

DÉDICACE

À mes grands-parents et à mon père, qui sont décédés. Leur mort aura été l'occasion première de philosopher sur la nature de mon existence, au même titre que ma naissance aura donné un sens à la leur.

TABLE DES MATIÈRES

REMERCIEMENTS.....	ii
DÉDICACE.....	iii
RÉSUMÉ ET MOTS-CLÉS.....	iv
INTRODUCTION.....	1
CHAPITRE I	
INTUITION D'ÊTRE UNE CHOSE PRÉSENTE À LA CONSCIENCE.....	7
1.1- Contexte historique.....	9
1.2- Réductionnisme de Parfit.....	12
1.2.1- Distinction réductionnisme et non-réductionnisme.....	12
1.2.2- Indétermination.....	15
1.2.3- Rôle du réductionnisme dans le contexte.....	17
1.2.4- Argumentation de D. Parfit.....	19
1.3- Objections substantialistes.....	26
1.4- Présence d'une chose dans le flux de conscience.....	32
1.4.1- Le corps.....	33
1.4.2- La cause de nos processus mentaux.....	34
1.4.3- L'interruption du potentiel de conscience.....	37
1.4.4-Révisionnisme ontologique.....	39

CHAPITRE II

AUTORITÉ À LA PREMIÈRE PERSONNE.....	43
2.1- Dépersonnalisation de la mémoire.....	44
2.1.1- Problème de circularité.....	44
2.1.2- Thèse de la description complète impersonnelle.....	44
2.1.3- Arguments.....	45
2.1.4- Implications sur la nature du moi.....	47
2.2- La mémoire comme source d'informations fiables et spontanés.....	48
2.2.1- Immunité à l'erreur d'identification.....	49
2.2.2- Illusion ou hypothèse confirmée de l'extérieur.....	50
2.2.3- Ensemble de présomptions acquises permettant l'action.....	51
2.2.4- Moi comme immersion dans un contexte autorisant un type de connaissance.....	51
2.3- Superfluité du moi dans la mémoire.....	54
2.3.1- Caractéristique non-spécifique à la 1 ^{ère} personne.....	55
2.3.2- Isoler l'identité du contexte autorisant des représentations spontanées.....	56
2.3.3- Support aux quasi-mémoires.....	61
2.4- Présence de l'illusion cartésienne.....	62
2.4.1- Autorité à la première personne et non-réductionnisme.....	63

2.4.2-	Dimension contre-intuitive du réductionnisme et de ses critiques.....	65
2.4.3-	Dissolution de la séparation radicale entre deux perspectives.....	66
2.5-	Résumé.....	68
CHAPITRE III		
	LE CONTEXTE DE NARRATION.....	70
3.1-	Approche de D. Parfit.....	72
3.1.1-	Morcellement.....	73
3.1.2-	Passivité.....	76
3.1.3-	Brièveté.....	77
3.2-	Le moi narrateur.....	77
3.2.1-	M. Schechtman.....	78
3.2.2-	A. Rudd.....	80
3.2.3-	Dimension active.....	80
3.2.4-	Application à la fission et aux séquences désignées par « ceci ».....	82
3.2.5-	Résumé.....	83
3.3-	Réponse aux objections centrées sur le contexte narratif.....	84
3.3.1-	Données empiriques.....	84
3.3.2-	Conditions non-nécessaires et insuffisantes.....	86
3.3.3-	Pronoms démonstratifs alternatifs.....	87

3.4-	Synthèse.....	89
3.4.1-	Observations courantes.....	90
3.4.2-	Les intuitions sous-jacentes au postulat d'une unité de signification.....	94
3.5-	Résumé.....	97

CHAPITRE IV

	UNITÉ DE L'EXPÉRIENCE DE SOI EN UN INSTANT.....	99
4.1-	L'argumentation de Parfit.....	101
4.2-	Possibilités empiriques d'une division de l'expérience consciente.....	103
4.2.1-	Les cerveaux divisés.....	103
4.2.2-	Critiques de leur utilisation.....	104
4.2.3-	Support à leur utilisation.....	104
4.2.4-	Synthèse de la discussion.....	106
4.2.5-	Résumé.....	110
4.3-	Possibilité conceptuelle d'une division de l'expérience consciente.....	111
4.3.1-	Moi défini comme entité à double face, publique et privée.....	111
4.3.2-	Autoréférence à la première personne occasionnellement publique et divisible.....	114
4.3.3-	L'attribution à soi comme intuition théorique confuse.....	117
4.4-	Résumé.....	123

CHAPITRE V

L'IMPORTANCE.....	126
5.1- « L'identité n'est pas ce qui importe dans la survivance ».....	128
5.1.1- Confusion inhérente au problème de l'identité du moi.....	128
5.1.2- Le paradoxe de la duplication.....	130
5.1.3- La solution de D. Parfit.....	132
5.1.4-Implications.....	133
5.1.5-Résumé.....	135
5.2- Les enjeux soulevés.....	135
5.2.1- Le tout importe-t-il plus que les parties ?.....	136
5.2.2- Exception à une règle ?.....	143
5.2.3-Récapitulation.....	146
5.3- Les représentations métaphysiques de la valeur de l'existence.....	147
5.3.1- La réalité naturaliste.....	147
5.3.2- La part de vérité et d'illusion essentialiste.....	149
5.4- Conclusion.....	155
CONCLUSION.....	156
BIBLIOGRAPHIE.....	161

RÉSUMÉ ET MOTS-CLÉS

Cette réflexion porte sur les illusions qui accompagnent la conscience de soi tel qu'elles se manifestent dans des questionnements sur l'identité du moi à travers le temps. Elle prend comme données les écrits de D. Parfit et des commentaires qu'ils ont suscités dans la littérature philosophique anglo-saxonne. Cinq aspects de la conscience de soi seront abordés : l'intuition d'y voir intérieurement une chose, l'autorité privilégiée que nous avons sur notre passé, l'intégration des événements ponctuels dans un récit de vie, l'unité phénoménale en un instant et l'importance de l'unité du moi à travers le temps. Ce travail questionne la présence en nous d'une intuition trompeuse selon laquelle nous sommes une chose de plus que le corps et un ensemble de processus mentaux particuliers, c'est-à-dire une croyance non réductionniste. Les arguments de D. Parfit en faveur d'une telle disposition à l'erreur, ainsi que la discussion qu'ils ont suscitée seront clarifiés de manière à évaluer si les comportements philosophiques sous-jacents en manifestent la présence. Effectivement, il s'y trouve une tendance à postuler arbitrairement certaines nécessités relatives à l'autoréférence et cette réaction est difficile à expliquer en l'absence d'une croyance non réductionniste.

Mots-clés : Identité personnelle, identité du moi à travers le temps, conscience de soi, autoréférence, non réductionnisme, réductionnisme, critère de la continuité physique et psychologique, indétermination, description complète impersonnelle, introspection, autorité privilégiée, moi narratif, unité phénoménale, importance et valeur de l'existence.

INTRODUCTION

Quel genre de chose suis-je ? Une âme ? Un corps ? Un intérieur capable de se réfléchir dans une pensée ? Une fiction ? Un mystère fabuleux et indéfinissable ? Une somme d'expériences reliées ensemble par des souvenirs, des croyances et des projets en continuité ?

Dans les traditions de R. Descartes, de J. Locke et de D. Hume, les philosophes anglo-saxons contemporains qui se sont posé cette question en appellent à l'imagination pour y répondre. Des fictions menaçant notre existence nous sont proposées. Notre cerveau est transplanté dans un autre corps. Notre corps est détruit, puis une copie parfaite en est ensuite créée de manière à reproduire un vécu indiscernable du nôtre, ni de l'intérieur, ni de l'extérieur. Une manipulation chirurgicale permet de créer deux êtres comme nous à partir de notre corps.

En envisageant ces scénarios, nous devrions identifier une limite au-delà de laquelle nous ne pourrions pas logiquement survivre. Elle serait l'essence de notre existence. Par exemple, si nous imaginons la réplication parfaite de notre matière comme une manière de mourir et d'être remplacé par un autre, alors l'exercice nous révèle que nous considérons notre cerveau et/ou notre corps actuel(s) comme étant ce à quoi réfère « je ». Si nous y voyons plutôt un moyen de changer de place, nous nous en faisons alors une idée différente.

De nombreuses distinctions philosophiques sophistiquées se sont construites autour de ces exercices. La principale range les théories en deux clans. D'un côté, il y a celles pour qui notre existence dépend essentiellement de la continuité d'un objet physique, que ce soit le corps ou le cerveau. De l'autre, elle dépend de la poursuite d'états mentaux comme des mémoires, des intentions, des traits de caractère et des croyances.

D. Parfit oppose plutôt le réductionnisme et le non réductionnisme. Tous les réductionnismes s'accordent sur une liste finie d'entités « familières » dont l'existence ne pose pas problème : le corps, le cerveau, les expériences, les pensées, les souvenirs, la poursuite d'intentions, la persistance de traits de caractère, leur enchaînement graduel dans le temps et un langage pour en parler. Elles se distinguent à propos des éléments constituant notre essence véritable. Certains diront que c'est une partie suffisante du cerveau. D'autres diront que c'est la poursuite des expériences peu importe le support physique. Cependant, ils ont en commun de nier l'existence d'une chose qui transcende cette liste finie d'entités. Le non réductionnisme pose, au contraire, que notre être ne peut pas être analysé à partir de ces éléments, mais qu'il renvoie à une entité supplémentaire. Il s'appréhende directement, sans passer par l'intermédiaire de composants plus fondamentaux¹.

Peu d'auteurs s'affichent comme non réductionnistes. D. Parfit amène cette distinction parce qu'il nous attribue une pulsion plus ou moins consciente à y croire². Le « nous » dont il parle s'observe dans le comportement des philosophes qui traitent de cette question. Ceux-ci nient se référer à une telle réalité transcendante, mais la manière dont ils réagissent aux divers exercices d'imagination semble montrer le contraire.

Prenons l'exemple suivant. Un philosophe X affirme qu'une table est une planche surélevée supportée par des objets verticaux et qui sert à manger. Il est donc réductionniste quant à la table, car il fait reposer son existence sur une liste finie d'entités plus fondamentales. Il y a devant X une planche à trois pieds du sol tenue par 4 morceaux de bois. Il mange dessus. Pourtant, il en est encore à se demander

¹ D. Parfit (1984), p.210-213.

² "In our thoughts about our identity, we are prone to illusions. That is why the so-called "problem cases" seem to raise problems: why we find it hard to believe that, when we know the other facts, it is an empty or a merely verbal question whether we shall still exist. Even if we accept Reductionist view, we may continue, at some level, to think and feel as if that view were not true." D. Parfit (1995), p.45.

s'il s'agit vraiment d'une table. Ce comportement philosophique étrange nous montre qu'il n'est pas réductionniste, bien qu'il prétende le contraire. Effectivement, s'il cherche encore, c'est parce qu'il croit que la table peut être une entité au-delà de ce qu'il a devant les yeux.

D. Parfit intervient dans le débat entourant la nature de notre existence en attribuant à ses collègues une telle méprise.

Premièrement, ils ne reconnaissent que l'existence du corps, du cerveau, d'expériences, puis d'événements mentaux qui les relient ensemble. Pourtant, ils s'opposent comme s'il y avait un fait de plus à propos duquel ils étaient en désaccord. Si vraiment ils étaient réductionnistes, ils reconnaîtraient que ce qui les oppose n'est pas réel, mais se réduit à choisir une perspective. Les uns voient dans le corps notre présence et les autres le font dans la poursuite d'événements mentaux. Les premiers refuseraient la réplication. Les seconds l'accepteraient. Plutôt que de s'opposer, ils ne s'attribuent simplement pas les mêmes éléments. Or, au lieu de tirer cette conclusion, ils avancent de nouveaux exercices d'introspection pour montrer aux autres qu'ils ne pointent pas à la bonne place lorsqu'ils disent « je ». Autrement dit, ils prétendent avoir une liste commune d'ingrédients devant eux pour nous définir, mais ils débattent comme si la réponse se trouvait ailleurs. Ils sont donc non réductionnistes à un niveau plus ou moins conscient. D. Parfit défend cette thèse en affirmant que si le réductionnisme est vrai, ces débats sont indéterminés, c'est-à-dire qu'il n'y a pas de « fait en la matière » censé trancher. Le nier suppose une croyance non réductionniste plus ou moins consciente³.

Deuxièmement, nombreux sont les auteurs qui affirment qu'un aspect de la vie mentale doit absolument renvoyer à notre existence pour être décrite adéquatement. Cette attitude est spécifique à nous et ne s'applique pas aux autres entités. Décrire

³ D. Parfit (1984), p.213-214. Le contenu ce paragraphe est une interprétation abrégée de sa théorie et je la défendrai tout au long de ce mémoire.

une rivière comme un objet qui passe à un endroit ou plutôt comme un flux continuels d'eau est facilement compris comme deux variantes d'une même réalité. Se référer seulement à la seconde description sans mentionner l'existence des rivières n'omet pas la présence d'une dimension fondamentale de ce qui s'y trouve. Autrement dit, la manière dont nous découpons les entités impliquées là est flexible. Cette affirmation est banale lorsqu'appliquée aux rivières. Cependant, lorsque les philosophes discutent de nos états mentaux, ils réagissent différemment. Ils refusent catégoriquement d'impliquer d'autres entités à la place de notre existence pour décrire la même réalité⁴. Ce comportement peut aussi trahir la présence d'une croyance non réductionniste. En effet, notre existence doit posséder quelque chose de plus qui n'est pas contenu dans la liste des entités reconnues par le réductionnisme pour l'expliquer. C'est pourquoi D. Parfit argumente en montrant comment l'unité de la vie et de la conscience peut être décrite adéquatement sans se référer à soi et donc de façon impersonnelle. Il emploie cette thèse notamment comme façon de séparer clairement les non-réductionnistes confus des réductionnistes assumés⁵.

La réflexion de D. Parfit autour de la nature de notre existence s'articule donc autour de l'affirmation selon laquelle la conscience que nous avons de nous-mêmes est profondément illusoire. Le comportement des philosophes, en supposant qu'il est révélateur de ce que nous pensons, est la donnée qui nous permet de le penser. Nous serions portés à voir dans les exercices d'imagination une chose mystérieuse qui échappe à toute analyse et à se duper nous-mêmes en prétendant le contraire. Cette tendance nous induirait en erreur et produirait des faux débats philosophiques. C'est la thèse centrale de ce mémoire.

⁴ D. Parfit (1999) élabore cette idée en détail.

⁵ Il affirme que le réductionnisme est difficile à saisir et que les thèses de l'indétermination et de la description complète impersonnelle visent à découvrir si nous sommes réductionnistes puisqu'il est difficile de le savoir. D. Parfit (1984), p.213, (1995), p.26-28 et (1999), p.221.

Ce mémoire clarifie les positions de D. Parfit et les enjeux qu'il soulève en passant en revue les auteurs qui l'ont commenté. Il en ressort que cette théorie de l'erreur est rarement prise en considération dans cette discussion. Les critiques de D. Parfit nient la présence d'une indétermination ou indiquent une dimension de la vie impossible à décrire sans se référer à soi. Ils prétendent rendre compte des intuitions du sens commun sans faire intervenir une chose qui transcende le corps, le cerveau et un ensemble de processus mentaux « familiers ». La discussion autour de cette démarche montre qu'en l'absence de cette réalité supplémentaire, D. Parfit a raison. J'y défendrai que de postuler la présence d'une intuition illusoire non reconnue par ces philosophes nous permet d'expliquer la résistance à ses arguments.

Ma recherche se divise en cinq parties correspondant chacune à un argument adressé contre la théorie de D. Parfit. J'y opposerai les pours et les contres avancés dans la littérature, pour y répondre ensuite.

Mon premier chapitre portera sur l'intuition que l'expérience consciente s'accompagne de la présence d'une chose singulière à laquelle elle doit être attribuée suite à une observation intérieure.

Dans le second chapitre, je traiterai le point de vue de la conscience de soi à travers le temps comme possédant une autorité distinctive, c'est-à-dire une façon d'user d'informations à propos du passé spontanément sans avoir à se justifier davantage.

Dans le troisième chapitre, je traiterai de l'intégration des expériences conscientes dans un récit de vie. Tous les états mentaux d'une personne prennent leur signification de la référence à celui-ci.

Dans le quatrième chapitre, je traiterai de l'unité phénoménale des expériences. Celle-ci peut être décrite comme le fait qu'il y a un effet que cela fait *pour moi* de vivre plusieurs expériences en même temps comme regarder cet écran et toucher à ce

clavier. Cependant, il n'y a pas d'effet que cela fait pour nous de le faire si je tape sur le clavier et que quelqu'un d'autre regarde l'écran⁶.

Dans le cinquième chapitre, je traiterai de l'importance spéciale qui accompagne l'identité du moi à travers le temps, plus précisément dans le cadre de l'unité de la conscience. L'identité entre moi maintenant et moi dans le futur correspond à ma survivance et un ensemble d'émotions, de pensées et de comportements spécifiques l'accompagnant. Cette idée est censée justifier une rationalité relative à soi distincte d'une rationalité relative aux autres.

⁶ Pour cette définition, voir T. Bayne et D. Chalmers (2003).

CHAPITRE I

L'INTUITION D'ÊTRE UNE CHOSE PRÉSENTE À LA CONSCIENCE

J. Locke définissait l'identité du moi à travers le temps par la faculté de se « consulter soi-même comme le même, comme une même chose qui pense en différents temps et en différents lieux »⁷. Cette remarque l'a amené à dissocier l'identité du moi de l'identité d'une substance, que ce soit une âme ou une chose matérielle. En effet, plusieurs âmes ou plusieurs corps peuvent logiquement incarner une seule et même conscience qui se représente comme étant la même à différents temps et à différents lieux. Inversement, une même âme ou un même corps peut ne pas se réfléchir comme étant le même à différents instants. Cette réflexion en a inspiré plusieurs, qualifiés de « néo-lockéens », à concevoir l'identité du moi comme continuité psychologique. Cette conclusion n'est pas consensuelle. D'autres croient encore que le moi consiste en une substance. Peu croit en l'existence d'une chose purement pensante. Ils préfèrent analyser le moi comme une chose physique munie d'une capacité mentale, mais qui n'est pas obligée de garder à la conscience son histoire pour être le même.

C'est dans ce contexte qu'intervient D. Parfit. Selon lui, il n'y a aucun fait qui oppose ces deux positions. Il s'agit d'une querelle verbale. Cependant, il estime que cette réponse s'oppose à une intuition implicite et fausse selon quoi nous sommes une chose de plus que le corps, le cerveau, ainsi que les pensées et les expériences particulières qu'ils causent. C'est la présence de cette illusion non reconnue qui explique la tournure insatisfaisante du débat.

Dans ce qui suit, je clarifierai cette idée en la situant d'abord dans son contexte. J'explicitai alors les principaux concepts et arguments avancés par D. Parfit en montrant comment ils vont dans cette direction. Je lui opposerai alors des arguments

⁷ J. Locke (2009), p.522.

visant à montrer que le moi est réellement un corps ou un cerveau particulier et que les approches héritées de J. Locke sont erronées. Je montrerai finalement comment la position de D. Parfit peut être défendue plus efficacement face à ces critiques en précisant la nature de l'illusion censée nous induire en erreur dans ce débat. Il s'agit de l'intuition selon quoi nous sommes une chose présente à la conscience.

1.1- Contexte historique

La réflexion de D. Parfit prend sa source dans celles de R. Descartes, de J. Locke et de D. Hume.

R. Descartes cherche à nous révéler la nature du moi au moyen d'une fiction. Je peux douter qu'un malin-génie me trompe sur l'existence de mon corps, mais je ne peux pas douter de mon existence, car ma pensée me révèle que j'existe. Il en conclut que je suis essentiellement une chose pensante et non une chose étendue dans l'espace⁸.

J. Locke considère aussi que je suis essentiellement pensant, mais conteste que je sois une substance sur la base de fictions différentes. Je peux m'imaginer être conscient dans le corps d'un valet, puis m'imaginer ensuite être conscient dans le corps d'un prince. Je peux m'imaginer changer d'âmes⁹. Il en conclut que je suis une capacité d'être conscient de moi-même. Si elle se transmettait d'une substance à une autre, je survivrais à cette succession, donc je ne suis pas une substance¹⁰.

D. Hume conçoit le moi comme une illusion d'unité à travers le temps. Il s'appuie sur la conscience qu'il a de ses perceptions. Il se cherche en elles et ne trouve que de multiples perceptions particulières distinctes de lui. Celles-ci sont changeantes. Ma perception d'aujourd'hui n'est pas celle d'hier, ni de demain¹¹. L'impression fausse

⁸ R. Descartes (1979), première, deuxième et sixième méditation.

⁹ J. Locke (2009), p.528-9.

¹⁰ « Le soi est cette chose pensante, intérieurement convaincue de ses propres actions (de quelque substance qu'elle soit formée, soit spirituelle, soit matérielle, simple ou composée, il n'importe) qui sent du plaisir et de la douleur, qui est capable de bonheur et de misère, et qui par là est intéressé pour soi-même aussi loin que cette conscience peut s'étendre. », J. Locke (2009), p.531.

¹¹ « Il y a certains philosophes qui imaginent que nous sommes à tout moment conscients de ce que nous appelons notre MOI, que nous sentons son existence et sa continuité d'existence, et que nous en sommes certains au-delà de l'évidence de la démonstration, aussi bien de sa parfaite identité que de sa parfaite simplicité. (...) Pour ma part, lorsque j'entre le plus intimement dans ce que j'appelle *moi-même*, je bute toujours sur quelque perception particulière ou sur une autre, de chaud ou de froid, de lumière ou d'ombre, d'amour, de haine, de douleur ou de plaisir. Je ne peux jamais, à aucun moment, me saisir *moi-même* sans une perception, et je ne puis observer autre chose que de la perception. » D. Hume, *Traité de la Nature humaine*, partie VI.

d'être une seule et même chose à travers le temps est produite par les relations de similarité, de contiguïté et de causalité entre nos perceptions¹².

Ils définissent ce que nous sommes en observant leurs pensées et sensations conscientes. La fiction chez R. Descartes et J. Locke constitue une introspection comparable à l'exercice de D. Hume. L'intuition sous-jacente, c'est que la présence de soi est indiquée par cette capacité d'auto-observation. Ils en tirent cependant des conclusions opposées. R. Descartes y voit le signe que nous sommes une chose pensante, J. Locke que nous ne sommes pas une chose, mais une capacité d'autoréflexion et D. Hume que nous sommes une fiction.

R. Swinburne adopte une position dans la continuité de Descartes en réaction à J. Locke et à D. Hume. Il défend que le moi est une chose mentale simple, indivisible, inanalysable dont l'existence ne peut jamais dépendre d'une convention¹³. Il argumente lui aussi à partir de fictions. Je peux m'imaginer changer de corps, changer de personnalité et survivre à une destruction complète de la conscience que j'ai de moi-même et exister sans corps. De plus, je peux imaginer une psychologie et un corps se divisant en deux, mais non mon être¹⁴. Je ne suis donc aucune de ces choses. Ensuite, contrairement à D. Hume, il fait l'expérience de lui-même dans l'unité des expériences en un instant, ainsi que dans leur brève succession. En ce moment, la vue de ce paragraphe s'accompagne de l'audition d'un bruit de voitures, puis de l'apparition d'une sensation de picotement. Il y a un effet que cela fait de vivre les trois ensembles. Selon R. Swinburne, cela témoigne de ma présence¹⁵.

¹² D. Hume, *Traité de la Nature humaine*, partie VI.

¹³ « The simple view claims explicitly that personal identity is one thing, and the extent of similarity in matter and apparent memory another. There is no contradiction in supposing that the one should occur without the other. (...) The simple view is normally combined with the view that person are indivisible. » R. Swinburne (1984), p.21-22.

¹⁴ R. Swinburne (1984), p.21-25.

¹⁵ R. Swinburne (1984), p.42-44 et p.47. J'ai modifié l'exemple. Il parle plutôt de l'expérience d'un mouvement, ainsi que de l'expérience d'une crampe combinée à la vision d'un flash bleu.

B. Williams affirme que nous sommes une chose physique : un corps. Considérer, comme J. Locke, que le moi est une capacité de se réfléchir à différents temps et différents lieux a l'inconvénient que plus d'une personne puissent la poursuivre. Si cette conscience de soi est transférable d'un corps à l'autre, elle pourrait l'être aussi dans deux corps distincts. Deux personnes en même temps seraient moi, ce qui semble impossible¹⁶. D'autre part, face à la menace de torture par un scientifique fou, la disparition de nos souvenirs ne nous empêcherait pas d'anticiper la douleur comme étant quelque chose qui nous arrivera à nous et non à quelqu'un d'autre¹⁷. B. Williams abandonne donc l'idée que le moi se caractérise par une capacité introspective, car, même si je la perds relativement à mon passé ou que j'anticipe de la perdre envers mon présent, je ne cesse pas d'exister. Cependant, il établit paradoxalement cette conclusion au moyen de l'introspection en imaginant des scénarios de duplication ou de lavage de cerveau. Il suppose qu'une forme d'observation intérieure me révèle que je suis un corps.

S. Shoemaker s'inspire de J. Locke en affirmant que nous sommes une chose physique distincte du corps dont les critères de survie sont psychologiques¹⁸. Il compare un morceau d'or et une statue. Il s'agit de deux choses distinctes, car elles n'ont pas les mêmes conditions de persistance. La statue peut être détruite, alors que le morceau d'or est conservé ou vice-versa. Pourtant, les deux choses coïncident dans la même substance et sont physiques. Il considère ainsi que c'est la continuation de ma psychologie qui est ma condition de survivance¹⁹. Je ne suis donc pas mon corps, car, si mon cerveau était transplanté dans un autre corps ou si ma psychologie était continuée par un autre moyen que mon corps actuel (par exemple, sa réplique

¹⁶ B. Williams (1973), p.20.

¹⁷ B. Williams (1970)

¹⁸ Centrée sur des états mentaux comme la mémoire, la poursuite d'intention, etc.

¹⁹ Locke se concentrait sur la conscience réfléchie de soi interprétée ensuite comme la mémoire des expériences passées, mais S. Shoemaker fait référence à un concept plus large de continuité psychologique qui inclue toutes formes d'états mentaux, même non-conscients. S. Shoemaker (1984), p.86-88.

cellules par cellules), je survivrais²⁰. Afin d'éviter le problème de la duplication mentionnée par B. Williams, il ajoute une clause selon quoi la continuation de la psychologie d'une personne ne doit pas s'embrancher²¹. Cette approche nie que je sois essentiellement une chose singulière, *une* substance, car je pourrais survivre au moyen de plusieurs corps si leur succession poursuivait les mêmes capacités psychologiques, conscience comprise²².

Ces trois auteurs contemporains partagent avec leurs prédécesseurs l'analyse du moi à partir d'une observation de l'expérience consciente, soit par l'imagination, soit par l'analyse de l'unité des expériences en un instant. Ce mode d'autoréflexion est donc une constante. Les divergences portent sur ce qui en découle et ce, principalement sur trois axes : physique ou mental, substantiel ou non, réel ou illusoire.

1.2- Le réductionnisme de Parfit

1.2.1- Distinction réductionnisme et non-réductionnisme

D. Parfit défend une position « réductionniste » qu'il définit ainsi : « A person's existence just consists in the existence of a brain and body, and the occurrence of a series of interrelated physical and mental events²³. » Les relations entre événements mentaux dont il parle sont les liens causaux entre un souvenir et une expérience, entre une intention et une action ou la persistance de croyances et de traits de caractères, ainsi qu'un enchaînement suffisamment important de telles relations à travers le

²⁰ S. Shoemaker (1984), p.89-90 et p.108-111.

²¹ S. Shoemaker (1984), p.85.

²² J'entends ici par « substance » l'idée d'une chose singulière définie par l'organisation de sa matière (physique ou non), par opposition à une chose composée de plusieurs unités de ce genre, mais définie plutôt par une apparence, une forme, une fonction, une visée commune. Si je survivais par le biais d'une série de réplique de corps, j'existerais alors clairement dans le sens d'un composite et non d'une matière singulière. Cette idée poursuit celle de J. Locke. S. Shoemaker (1997) élabore des nuances similaires.

²³ D. Parfit (1984), p.211.

temps²⁴. De « mental », il n'offre aucune définition et reste neutre quant à la question du dualisme ou du monisme²⁵.

Parfit analyse le moi comme un composé d'éléments plus particuliers. Il en compare l'existence à la relation unissant une multitude d'arbres à une forêt ou une diversité de personnes à une nation²⁶. Il qualifie cette relation de deux manières.

Premièrement, il peut s'agir d'une identité. La forêt peut être la multiplicité d'arbres comme la nation peut être la diversité de personnes. Dans ce cas, le moi est le corps, le cerveau et/ou un ensemble de pensées et d'expériences reliées. Il ne privilégie cependant pas cette relation pour caractériser le moi, même s'il la range parmi les théories réductionnistes. « Some Reductionists claim (4) a Person *just is* a particular brain and body, and such a series of interrelated event²⁷. »

Deuxièmement, il peut s'agir d'exister « distinctement, mais non séparément » de ces éléments. « Other Reductionists claim (5) a Person is an entity that *is distinct* from a brain and body, and such series of events²⁸. » D. Parfit reprend ainsi l'idiome de la cohabitation avancé par S. Shoemaker. Deux entités peuvent occuper la même portion de la réalité sous le modèle d'un morceau de marbre et d'une statue²⁹. Il y a deux entités *distinctes* et non pas une, car certains énoncés peuvent être vrais d'une entité et non de l'autre. La statue peut être graduellement constituée d'une nouvelle matière et non le morceau de marbre. De la même manière, le cerveau ou un ensemble d'états mentaux ne peuvent pas penser, alors qu'un moi le peut. C'est en ce sens qu'ils sont distincts³⁰. Cependant, ils n'existent pas « séparément » au sens où

²⁴ D. Parfit (1984), p.205-206.

²⁵ D. Parfit (1984), p.241.

²⁶ D. Parfit (1984) p.210 et p.212 et (1995), p.19-20.

²⁷ D. Parfit, p.211.

²⁸ D. Parfit, p.211.

²⁹ D. Parfit (1995), p.16-17.

³⁰ Cette signification du caractère distinct mais non séparé du moi apparaît clairement lorsque D. Parfit précise l'analogie avec la nation : « When we use the word « France » to refer to a nation, we are not referring to something other than a nation. We are not referring to this nation's government, or its

c'est la même substance qui est pointée du doigt en désignant les deux entités. « Since we are not separately existing entities, we would not need not to be separately listed in an inventory of what exists³¹. »

D. Parfit précise cette « non-séparation » de deux façons. D'une part, il y a une entité qui est à la base de l'autre, c'est-à-dire que connaître ce qui lui arrive suffit à établir ce qui advient de l'autre. Si nous savons qu'il y a un corps, un cerveau et un ensemble d'états mentaux, nous avons donc tous les ingrédients pour déterminer ce qui arrive de moi³². De la même manière, si nous savons qu'il y a un morceau de marbre configuré d'une certaine manière, nous avons tous les outils pour conclure sur ce qui advient de la statue. D'autre part, relativement à un corps, à un cerveau et à un ensemble d'événements mentaux inter-reliés, l'existence du moi n'ajoute qu'un langage de plus pour parler de la même réalité et non un fait qui la transcende³³.

Cette position sert à analyser la nature de mon existence et l'identité numérique du moi à travers le temps. Cette dernière relation renvoie au nombre d'êtres impliqués dans une situation et donc au fait qu'un être en un instant et un être en un autre instant sont un seul et même être. Certains diront que la persistance d'un objet physique détermine cette identité, alors que d'autres diront que c'est la poursuite de représentations mentales, de projets, de traits de caractère et de croyances³⁴. Le trait commun de toutes les approches réductionnistes est de préciser ces concepts en référant à d'autres entités et non au moi lui-même. « They are Reductionist because

citizens, or to its territory. This can be shown as follows. If "France" referred to the French government, France would cease to exist if the government resigned and there was a period of anarchy. But this is false. (...) And if "France" referred to these citizens, it must have been these citizens that declare war on Germany. This is also false.", D. Parfit (1984), p.472.

³¹ D. Parfit (1999), p.221.

³² D. Parfit (1984), p.213.

³³ D. Parfit (1984): "(...) on our concept of a person, people are not thoughts and acts. They are thinkers and agents. I am not a series of experiences, but the person who has these experiences. A Reductionist can admit that, in this sense, a person is what has experiences, or the subject of experiences. This is true because of the way we talk." p.223. Il élabore ce point dans D. Parfit (1995), p.19-28 et D. Parfit (1999).

³⁴ D. Parfit (1984), p.202-209.

they claim that the fact of a person's identity over time just consists in the holding of certain more particular facts³⁵. »

Il oppose cette conception « réductionniste » à ce qu'il qualifie de « non réductionnisme », qui pose que le moi est une chose singulière de plus qu'un tel ensemble de parties. « On this view, personal identity over time does not consist in physical and/or psychological continuity. It involves a further fact³⁶. » Les positions de R. Descartes et R. Swinburne en sont des exemples.

1.2.2- Indétermination

Afin de clarifier cette opposition, il affirme que nous pouvons imaginer des situations où notre existence et notre identité à travers le temps sont indéterminées. C'est-à-dire que nous avons une description complète de la situation et notre existence se ramène à un choix concernant la manière d'y préciser les contours flous du moi³⁷. Cette thèse sert de test pour départager les réductionnistes des non-réductionnistes³⁸.

La méthode employée pour opposer les diverses approches à propos du genre de chose que le moi est, c'est souvent d'imaginer des scénarios où des composants habituellement associés à mon existence sont séparés en imagination ; puis de demander ce qui advient de mon existence afin de cerner ceux qui lui sont essentiels et ceux qui ne lui sont que contingents. Le scénario du malin-génie de R. Descartes sépare en imagination notre activité de penser de l'action d'une chose étendue pour en conclure que nous sommes essentiellement une chose pensante. J. Locke sépare

³⁵ D. Parfit (1984) p.210.

³⁶ D. Parfit (1984), p.210.

³⁷ D. Parfit (1984), p.213-214.

³⁸ « My claim about Reductionist draw distinction that, in this abstract form, are hard to grasp. But there are other ways of discovering whether we are Reductionist in our view of some kind of thing. If we accept a Reductionist View, we shall believe that identity of such a thing me be, in quite unpuzzling way, indeterminate.» D. Parfit (1984), p.213.

en imagination la conscience de soi des substances la supportant pour en conclure que nous sommes essentiellement une conscience réfléchie et non une chose.

Une des conclusions que D. Parfit tire de cet exercice, c'est que nous y cherchons une réalité au-delà des composants que nous séparons en imagination.

Prenons la réplique. J'actionne une machine qui enregistre l'état de mon corps, le détruit, envoie l'information par ondes radios sur Mars et en crée à partir de nouveaux matériaux une copie parfaite. L'être qui en résulte poursuit mes projets, traits de caractère, croyances et souvenirs exactement de la manière que je l'aurais fait. La conscience qu'il a de moi est indiscernable de celle que j'ai habituellement de moi-même³⁹. Est-ce que je suis cet être ou est-ce que c'est quelqu'un d'autre ? Ce scénario sépare en imagination la continuation de ma psychologie et l'existence continue de mon corps afin de cerner laquelle des deux composantes est essentielle à mon existence⁴⁰. L'approche de J. Locke et de S. Shoemaker cible la continuation de ma psychologie, alors que l'approche de B. Williams cible mon corps.

D. Parfit devine plutôt en nous une insatisfaction qui nous pousse à chercher quelque chose de plus dans ce scénario. La description qui en est faite nous apparaît incomplète, comme si quelque chose manquait : moi. « When I fear that, in Teletransportation, I shall not get to Mars, my fear is that the abnormal cause may fail to produce this further fact⁴¹. » Il nomme cet élément un « fait supplémentaire profond » ou une « substance mentale simple » pour désigner une réalité ultime et inanalysable au-delà des composants présentés à nous en imagination⁴². En ce sens, une partie de nos intuitions vont dans le sens de R. Descartes et R. Swinburne⁴³.

³⁹ D. Parfit (1984), 199-200.

⁴⁰ D. Parfit (1995), Section I. D. Parfit (1995), p.202-209.

⁴¹ D. Parfit (1984), p.279.

⁴² D. Parfit (1995), p.242-243.

⁴³ "In our thoughts about our identity, we are prone to illusions. That is why the so-called "problem cases" seem to raise problems: why we find it hard to believe that, when we know the other facts, it is

Il présente son réductionnisme comme une famille d'approches qui nient cette intuition. Elles ont une ontologie commune. Elles ne reconnaissent comme entités réelles que le corps, le cerveau et un ensemble d'événements et de processus mentaux⁴⁴. Elles ne sont donc pas en désaccord à propos d'une réalité, mais à propos de la manière d'y préciser le concept « moi ». Elles impliquent donc qu'il n'y a rien de plus à chercher dans la réplication et que les deux réponses (je suis la réplique ou je ne le suis pas) ne sont qu'un choix de *gestalt* comparable à voir un verre à moitié vide ou à moitié plein⁴⁵. Autrement dit, j'ai en ma possession tous les éléments qui me permettent de comprendre ce qui se passe, mais la présence ou l'absence de « moi » dépend d'un choix concernant une situation à propos de laquelle aucune convention linguistique n'a été prévue⁴⁶. Une des deux réponses peut être préférable à l'autre pour diverses raisons, mais non pas simplement parce qu'elle pointe le « moi véritable »⁴⁷.

1.2.3- Rôle du réductionnisme dans le contexte

La réflexion de D. Parfit poursuit celle de D. Hume au sens où l'auto-observation nous révélerait la présence d'un moi illusoire. D. Hume employait l'introspection comme méthode pour découvrir son absence parmi les perceptions malgré une pulsion forte à vouloir se trouver là. D. Parfit fait la même chose⁴⁸. La réplication de soi nous révèle une attitude de fouille et d'inconfort. Nous réagissons comme D.

an empty or a merely verbal question whether we shall still exist. Even if we accept Reductionist view, we may continue, at some level, to think and feel as if that view were not true." D. Parfit (1995), p.45.

⁴⁴ « So while believers in the different criteria disagree about imaginary cases, they agree about what is in fact involved in the continued existence of most actual people ». D. Parfit (1984), p.211.

⁴⁵ D. Parfit (1999), p.259-261.

⁴⁶ D. Parfit (1984), p.213-214 et D. Parfit (1995), Section I.

⁴⁷ Par exemple, il peut être plus utile de considérer la réplique comme étant la même personne ou non.

⁴⁸ Il résume ainsi son usage des expériences de pensée: « What we can learn from this imaginary story ? Some believe that we can learn little. (...) This criticism might be justified if, considering such imagined cases, we had no reactions. But these cases arouse in most of us strong beliefs, not about our words, but about ourselves." D. Parfit (1984), p.200.

Hume. Nous cherchons parmi notre idée du corps ou de nos processus mentaux où notre existence se trouve. Va-t-elle avec l'un ou avec l'autre ?

La présence de cette entité fuyante est ce que nie le réductionnisme. Le moi consiste en un corps et en une série d'événements mentaux et physiques inter-reliés parce qu'au-delà de ces composantes et le langage pour en parler, il n'y a rien de plus à chercher pour se référer à nous-mêmes avec succès. Pourtant, nous avons, sans forcément nous en rendre compte au premier coup d'œil, une tendance à y chercher quelque chose de plus.

La méthode de D. Parfit est une extension de celle du malin-génie de R. Descartes, reprise par J. Locke et leurs successeurs. Elle nous place dans un état mental susceptible de nous faire découvrir la nature du moi de façon claire et distincte, par opposition au contexte vague de la réalité courante. Dans la confusion, nous pouvons prétendre être réductionniste⁴⁹. Face à des options claires, notre attitude montre le contraire. Parfit estime donc que Descartes saisit par ce moyen une intuition du sens commun⁵⁰. Il se distingue de lui en niant que nous pointons là une réalité.

Ensuite, cette position répond à l'opposition entre J. Locke, S. Shoemaker et B. Williams, c'est-à-dire les conceptions du moi centrées sur la continuation de processus mentaux ou sur l'existence d'une chose physique. Selon Parfit, ce débat se nourrit d'une croyance non réductionniste implicite. De nombreux auteurs y interviennent sans reconnaître la part d'illusion qu'ils y mettent⁵¹. Si le moi consiste

⁴⁹ « In ordinary cases, questions about our identity have answers. In such cases, there is a fact about personal identity, and Reductionism is one view about what kind of fact this is. (...) We may find it hard to decide whether we accept this view, since it may be far from clear when one fact just consist in another. We may even doubt whether Reductionists and their critics really disagree." D. Parfit (1995), p.27.

⁵⁰ D. Parfit (1995), p.27-28.

⁵¹ « (...) many of us, I suspect, have inconsistent beliefs. If we are asked whether we believe that there are Cartesian Egos, we may say No. And we may accept that, as Reductionists claim, the existence of a Person just involves the existence of a body, and the occurrence of a series of interrelated mental and physical events. But, as our reactions to the problem cases show, we don't fully accept that view." D. Parfit (1995), p.28.

en un corps et en une série d'événements mentaux inter-reliés, il n'y a là qu'un problème verbal, d'où l'appel à la notion d'indétermination. Que l'un proclame avoir trouvé le vrai moi dans le corps et que l'autre prétende l'avoir plutôt fait dans la série d'événements mentaux par le détour d'une imagerie mentale trahit chez ces auteurs la présence d'une contradiction. S'ils ne reconnaissent pas la présence d'une chose de plus qui transcende ces composantes, alors il n'y a aucun sens à affirmer qu'ils y ont trouvé quoi que ce soit, puis ensuite de multiplier les exercices pour faire voir à leurs adversaires réticents ce qu'ils viennent d'y découvrir. Pourtant, la littérature entourant l'identité du moi à travers le temps est remplie de tels comportements⁵².

La solution de Parfit consiste à dévoiler l'apport de cette illusion. Une fois clarifiés, les problèmes restants sont plus faciles à résoudre. Il reste alors à déterminer laquelle de ces composantes justifie nos attitudes envers le moi. Selon lui, elles reposent sur la continuation de nos processus mentaux, mais en aucun cas sur l'existence du corps⁵³. C'est par ce détour qu'il finit par supporter une approche comme celle de J. Locke à l'encontre d'approches comme celle de B. Williams⁵⁴.

1.2.4- Argumentation de D. Parfit

Les arguments de D. Parfit en faveur du réductionnisme suivent la méthode de D. Hume. Ils proposent des fictions comme outil d'introspection de manière à montrer de façon claire et distincte l'absence d'une réalité qui transcende le corps et les processus mentaux à laquelle nous renvoyons dans une expérience subjective.

⁵² D. Parfit *nous* attribue régulièrement des dispositions à réagir d'une certaine manière. Le rationnel qui le justifie apparaît largement provenir du comportement observable des philosophes qui se posent ces questions, en présumant que ceux-ci représentent « nous ».

⁵³ D. Parfit (1984), p.261-266, p.282-287, p.307-318. Cette question sera l'objet du cinquième chapitre.

⁵⁴ Cet aspect de la réflexion de Parfit sera discuté plus en détail dans le cinquième chapitre.

1.2.4.1- Les quasi-mémoires

La mémoire ne nous révèle pas la présence d'un moi transcendant le corps, les pensées et les expériences particulières. Pour l'illustrer, il avance le concept de quasi-mémoire. C'est une représentation dont l'apparence et la cause sont similaires à la mémoire, mais qui peut se transmettre entre deux êtres. Il imagine ainsi qu'une trace laissée dans le cerveau de Paul suite à son voyage à Venise est transplantée dans le cerveau de Jane. Cette dernière peut alors visualiser la scène vivement comme si elle s'en souvenait. La mémoire normale peut être conçue alors comme une somme plus grande de telles traces reliant ainsi une perspective présente à un passé indépendamment de la présence d'un moi⁵⁵.

Il nie alors que l'introspection nous révèle dans la mémoire la présence d'une réalité supplémentaire à cette notion de quasi-mémoire. Il le fait en se plaçant du point de vue de la réplique parfaite d'un moi discutée plus haut. Elle aurait les mêmes raisons que nous de croire qu'elle a vécu les expériences qu'elle se représente, mais elle aurait tort⁵⁶.

Les deux arguments montrent qu'analysée de l'intérieur, la mémoire ne nous révèle rien de plus qu'une continuité psychologique avec l'expérience passée et non la présence d'une chose pensante⁵⁷.

1.2.4.2- Les spectres

Parfit imagine des spectres d'opération allant de la destruction complète de mon corps suivie d'une réplication parfaite à une situation où rien ne se passe. Entre les deux, il y a des opérations où 45%, 46%, 47%, 50 %, 51% de mes cellules seront

⁵⁵ D. Parfit (1984), p.220-222. Cet argument sera analysé plus en détail au 2^{ème} et 3^{ème} chapitre.

⁵⁶ D. Parfit, p.223-224.

⁵⁷ « (...) it seem to show that we could not tell, from the content of our experiences, whether we really are aware of the continued existence of a separately existing subject of experiences. The most we have are states of mind like that of my Replica", D. Parfit (1984), p.224.

remplacées par des copies. Dans tous les cas, ma psychologie serait continuée de façon similaire. Si je considère être une chose physique (cerveau ou corps), je devrai alors admettre qu'au milieu du spectre, j'aurai une description complète de la situation, alors que mon existence ou non sera un choix concernant l'usage d'un mot⁵⁸. Il imagine ensuite un spectre d'opérations similaires où, dans le milieu, il s'agirait plutôt d'un mélange des cellules, des souvenirs, des traits de caractère, des intentions de moi et d'une autre personne pour étendre cet argument également aux approches de la continuité psychologique⁵⁹.

Ces spectres illustrent comment l'existence du corps, du cerveau et d'une série d'événements mentaux inter reliés peuvent être entièrement compris, alors que la présence du moi en eux n'ajoute rien de plus qu'un choix de langage⁶⁰.

1.2.4.3- Cogito révisé

D. Parfit reprend le cogito de Descartes afin de déduire l'existence du moi en partant d'étapes différentes. Si Descartes dit douter, il y a un doute qui survient et donc une pensée qui est en cours, mais il est précipité d'en conclure qu'une chose doute⁶¹. Plutôt que de faire intervenir une chose qui se positionne par rapport à la pensée, D. Parfit en appelle aux pronoms « ceci », « cette », « dans » et utilise le corps pour faire le même travail⁶². Au lieu de « je pense », il peut être dit « Cette pensée a lieu dans cette vie causée par ce corps ». À cette étape, qu'est-ce qui nous permettrait de conclure en l'existence d'un moi qui pense ? La structure verbale de la pensée et non l'observation directe d'une chose pensante. Comme les pensées sont

⁵⁸ D. Parfit (1984), p.234-235.

⁵⁹ D. Parfit (1984), p.236-241.

⁶⁰ D. Parfit (1984), p.241-242 et D. Parfit (1995), p.20-28. Cet argument sera approfondi au 3^{ème} chapitre.

⁶¹ D. Parfit (1999), p.233-237 et D. Parfit (1984), p.224-226.

⁶² D. Parfit (1984), p.226, p.252 et D. Parfit (1999), p.209.

grammaticalement attribuées à des penseurs, « Cette pensée a lieu dans cette vie causée par ce corps » nous autorise à en conclure : « je pense »⁶³.

1.2.4.4- Examen de physique

Un phénomène qui peut apparemment nous révéler la présence d'un moi à l'œuvre dans la succession des expériences conscientes, c'est l'unité phénoménale des expériences en un instant et dans leur brève succession. La sensation de taper sur un clavier, la pensée présente, la couleur des caractères et la musique en arrière-fond forment une seule et même expérience. Il y a UN effet que cela fait d'en faire l'expérience. R. Swinburne défendait que cela témoigne de la présence d'un moi simple et inanalysable dans le flux de conscience.

La réponse de Parfit ne consiste pas à nier la présence d'une unité d'expérience, mais à la traiter comme une réalité distincte de moi. Il imagine le scénario suivant. Je suis muni d'un engin qui divise mon flux de conscience afin d'effectuer une série de calculs avec mon hémisphère gauche et une autre série avec mon hémisphère droit. Pendant dix minutes, il y a un effet que cela fait de calculer avec l'hémisphère gauche et un effet que cela fait de calculer avec l'hémisphère droit, mais il n'y a pas d'effet que cela fait d'effectuer les deux en même temps. Dans un de mes flux de conscience, je me vois calculer dans l'autre comme je verrais un frère siamois le faire et vice-versa. Après dix minutes, je réunifie ma conscience et en retiens un souvenir vif d'avoir calculé deux séries d'exercices⁶⁴.

Dans la mesure où ce qui pense dans cette situation n'est pas une entité de 10 minutes occupant un hémisphère, mais plutôt une chose qui se trouve dans les deux simultanément, l'unité de chacune des expériences ne peut pas être décrite en référant à un moi. Effectivement, cela suggère faussement qu'il n'y a qu'une seule unité

⁶³ D. Parfit s'éloigne ici cependant du cogito de Descartes, car il reconnaît d'abord l'existence du corps. Paul Bernier (2010) p.597 fait cette analyse.

⁶⁴ D. Parfit (1984), p.245-248.

phénoménale, alors qu'il y en a deux. Parfit présente plutôt chacune des unités comme un fait brut. Autrement dit, ce scénario montre un cas où pointer l'unité des expériences n'équivaut pas à pointer la présence d'un sujet. Pour pointer alors un moi, il faut désigner la chose phénoménalement désunie et située dans les deux hémisphères. Constaté ainsi la présence d'une unité d'expérience n'équivaut pas à observer intérieurement la présence d'un moi⁶⁵.

1.2.4.5- Fission

Nous avons vu que B. Williams objectait à J. Locke que le moi ne pouvait pas être une conscience réfléchie indépendamment de la substance la supportant, car deux personnes pourraient avoir la conscience de ma vie et prétendre légitimement être moi, ce qui apparaît contradictoire. R. Swinburne a objecté que n'importe quel critère centré sur la continuité d'une chose physique contient le même problème. Un cerveau, par exemple, pourrait être divisé en deux avec une capacité égale de continuer ma psychologie et être transplanté dans deux corps exactement similaires à celui du moi d'origine. Si l'existence continue de la moitié de mon cerveau suffit à assurer mon existence, alors les deux personnes résultantes auraient des raisons égales et légitimes de prétendre être moi. R. Swinburne en conclut que nous ne sommes ni une conscience réfléchie, ni une chose physique, mais une chose de plus qui est mentale, simple, indivisible et inanalysable⁶⁶.

D. Parfit en déduit plutôt qu'il s'agit d'un cas d'indétermination. Nous pouvons étendre notre concept de « moi » de manière à conclure que nous serions après la division une seule et même personne dont l'esprit est à deux endroits en même temps. Il s'agirait d'une extension de l'examen de physique où je peux avoir deux flux de conscience séparés simultanément et communiquer avec moi-même par des moyens publics. Nous pouvons aussi étendre notre concept de « moi » de manière à conclure

⁶⁵ D. Parfit (1984), p.246-252. Cet argument sera analysé plus en détail dans le 4^{ème} chapitre.

⁶⁶ R. Swinburne (1984), p.13-17.

que je serais détruit et remplacé par deux nouvelles personnes⁶⁷. Or, même si nous ne statuons pas là-dessus, il est possible de comprendre entièrement ce qui se passe. La manière de préciser les contours du moi dans ce cas est un choix de mots⁶⁸.

Plusieurs éléments défient nos intuitions dans ce scénario. Premièrement, il nous apparaît étrange que deux parties d'un moi puisse être à des endroits différents et interagir entre elles à la manière de deux personnes, chacune avec leur histoire séparée⁶⁹. Deuxièmement, il nous apparaît étrange que la relation entre moi et un être futur puisse être intrinsèquement la même, alors qu'un événement externe à cette relation puisse déterminer si cet être sera moi ou un autre. S'il n'y a qu'une des deux opérations de transplantation d'hémisphère qui réussisse, je survivrai, mais s'il y en a deux, je mourrai. Pourtant, ma relation à l'être résultant sera la même. Seulement ma relation à un autre être viendra m'empêcher de lui appliquer le concept « moi »⁷⁰. En ce sens-là, il n'y a rien qui disparaît de l'univers hormis le droit d'user d'un mot et cela nous apparaît étrange comme façon de mourir⁷¹.

D. Parfit en conclut principalement que l'existence du « moi » n'est pas ce qui importe dans la survie⁷². Cette question sera traitée au 5ème chapitre. Cependant, ce scénario a aussi des implications relativement à la présence ou non d'une chose de plus que le corps, le cerveau et une série de pensées et d'expériences reliées dans la conscience de soi. Prenant pour acquis que la fission peut me tuer, je peux m'imaginer cesser d'exister alors que rien dans le flux de conscience ne l'indique. Mon existence dépend d'une relation qui lui est extrinsèque. Donc, l'unité de ce flux de conscience à travers le temps ne nécessite pas ma présence pour être comprise.

⁶⁷ D. Parfit (1984), p.256.

⁶⁸ D. Parfit (1984), p.258.

⁶⁹ D. Parfit (1984), p.256-257.

⁷⁰ D. Parfit (1984), p.261, 262, 267. Cet argument sera discuté tout au long de ce mémoire.

⁷¹ D. Parfit (1995), section III

⁷² D. Parfit (1984), p.263.

1.2.4.6- Synthèse

Selon Parfit, l'usage de telles fictions serait injustifié si les états mentaux étaient des modifications d'une substance autre que le corps dont l'indivisibilité est une caractéristique essentielle⁷³. En effet, il serait alors illégitime d'imaginer que des morceaux de cette substance soit transférés à une autre, divisés en deux et combinés. Il présuppose cependant que cette théorie est fausse et que les états mentaux dépendent de la modification d'une substance physique. Ce n'est pas l'introspection qui nous révèle ce fait, mais des connaissances empiriques⁷⁴. Ensuite, il fait la supposition que si l'esprit dépend du corps, alors rien n'empêche d'imaginer a priori qu'il puisse être divisé, morcelé, recomposé, répliqué de cette manière. Il ne précise pas pourquoi, mais nous pouvons présumer que c'est parce que les choses physiques, *en général*, le peuvent.

Ces arguments ont une structure commune. D. Parfit indique un composant censé nous révéler la présence d'un moi non réductionniste, puis il imagine ensuite une fiction nous montrant comment il n'y a rien de tel qui l'accompagne. Il pointe la mémoire, puis invoque un transplant de traces entre cerveaux ou une réplique pour nier qu'elle nous le révèle. Il pointe le corps, puis imagine un spectre de recombinaison partielle avec des cellules nouvelles pour nier qu'il y a en lui un moi qui le transcende. En effet, dans le milieu du spectre, il y a une telle substance physique, mais l'attribution à moi n'est que l'addition d'un langage. Suivant la même logique, il pointe la pensée en revisitant le cogito de Descartes, puis l'unité des expériences phénoménales en lui opposant l'examen de physique et finalement la continuité de nos processus physiques et mentaux à travers le temps en imaginant une fission. Autrement dit, il fait la même chose que D. Hume, mais en employant un imaginaire de science-fiction. Il cherche le moi au moyen d'un exercice

⁷³ D. Parfit (1984), p.238, p.259.

⁷⁴ D. Parfit (1984), p.227-228, p.245-248.

d'introspection et ne trouve qu'un mot pour parler de toutes les composantes qu'il appréhende et rien de plus.

1.3- Objections substantialistes

D. Parfit défend une position qu'il qualifie de réductionniste. Elle inclut un ensemble de thèses impliquant que l'existence du moi consiste en l'existence d'un corps et d'une série d'événements et de processus mentaux inter reliés et non une chose qui transcende ces composantes. Il avance que la possible indétermination de l'identité du moi à travers le temps sert à départager plus clairement cette approche du non réductionnisme. Ces précisions visent à solutionner un ensemble de problèmes discutés par R. Descartes, J. Locke, D. Hume et leurs héritiers contemporains. Elles tranchent le débat entre ceux qui présentent notre existence comme l'existence d'une chose physique singulière et ceux qui la définissent plutôt par une chose dont les conditions de survivance sont un ensemble de relations psychologiques. Selon lui, en l'absence d'une chose qui transcende les composantes actuelles du moi, cette opposition ne porte pas sur la réalité, mais sur la manière de la nommer. Or, les résistances de nombreux auteurs à tirer cette conclusion suggèrent la présence d'une croyance non réductionniste implicite. C'est ainsi qu'il cherche à défaire cette intuition au moyen de diverses stratégies d'imageries mentales.

Rares sont les auteurs contemporains selon qui nous sommes une chose pensante simple séparée du corps dans la tradition de Descartes. Cependant, ils sont plusieurs à considérer le moi comme une chose dont les frontières sont réelles et non pas définies par un ensemble d'événements psychologiques et physiques inter-reliés dont les contours peuvent être précisés de plus d'une manière. Cette section présentera des objections qui rejettent autant que Parfit l'existence de la substance cartésienne, mais qui considèrent le moi comme étant une chose au sens où ses frontières et sa présence sont indépendantes d'un choix langagier. Plutôt que de s'appuyer sur une méthode

introspective, certains auteurs estiment que des considérations d'économie ontologique ou une théorie causale de la référence permettent de fixer les frontières du moi dans celles du corps ou du cerveau. D'autres expliquent notre résistance à imaginer notre survivance par réplication en vertu d'une interruption du potentiel de conscience.

E. Olson, suivant B. Williams, défend une position qu'il qualifie d'animaliste. Nous sommes des organismes biologiques pensants. Ce qui fait notre unité est une structure physique d'auto-organisation⁷⁵. Notre psychologie est contingente. Des éléments comme la conscience que j'ai de moi-même à travers le temps ne déterminent pas mon existence. Si je la perds complètement pour repartir ensuite à neuf, je ne meurs pas tant que l'organisme qu'est mon corps continue de fonctionner. Une transplantation de cerveau ne me permettrait pas de changer de corps, car le cerveau n'est qu'un organe de l'organisme qui correspond au moi⁷⁶. Selon lui, c'est la conception du sens commun⁷⁷.

Son argumentation ne s'appuie pas sur des expériences de pensée, mais sur une considération d'économie ontologique⁷⁸. L'organisme pensant, l'animal, répond aux conditions pour être un moi. Il pense, il vit des expériences et il est capable de se réfléchir dans une pensée. Si nous admettons, comme S. Shoemaker, l'existence d'une autre entité qui coïncide matériellement avec lui, mais qui s'en distingue à la manière d'une statue par rapport à un morceau de marbre, nous nous retrouvons avec

⁷⁵ E. Olson (2007), p.27-29.

⁷⁶ E. Olson (2007), p.39-44.

⁷⁷ « No one is going to feel immediately drawn to any alternative views – that we are bundles of perception, or immaterial substances, or non-animals made of the same matter as animals, say. Compared with those proposals, the idea that we are animals looks like plain common sense. » E. Olson (2007), p.23.

⁷⁸ Cet aspect est largement élaboré dans E. Olson (2007), Chapitre 9, section 5 et 6.

deux moi à la même place sans avoir moyen de savoir auquel des deux le mot « je » fait référence⁷⁹.

D'autre part, un ensemble de processus mentaux ne pense pas, car ce n'est pas une chose. Si je perds complètement la mémoire, ce n'est donc pas une chose qui est détruite. La possibilité d'être détruit et répliqué implique d'être un système psychologique, alors qu'un système ne pense pas. Seule une chose pense. Il y a donc une confusion catégorielle de la part de ceux qui définissent le moi par la poursuite d'états mentaux⁸⁰. Dans le cadre d'une ontologie qui vise à limiter la multiplication inutile des entités du monde, cette conséquence est inacceptable. Les frontières de l'organisme biologique étant déterminées par une organisation physique concrète et indépendante de notre esprit, il est plus simple d'y tracer les frontières du moi⁸¹.

Ces considérations constituent des objections à Parfit, car elles nient qu'il y a indétermination entre une conception du moi comme chose physique et comme chose dont la survivance dépend d'une forme de continuité psychologique. Le moi ne consiste donc pas en un corps *et* un ensemble d'événements mentaux inter reliés, mais *seulement* dans un corps pensant. En ce sens, une copie physique de moi ne serait pas, comme le suggère D. Parfit, potentiellement moi ou un autre selon un choix de mots. Elle ne serait pas moi, point final. E. Olson répond que les frontières

⁷⁹ « Consider what it would mean if you were not the animal. The animal thinks. And of course you think. (...) So if you are not that thinking animal, there would be *two* beings thinking your thoughts: there would be the thinking animal, and there would be you, a thinking non-animal. We should each share our thoughts with an animal numerically different from us. (...) In any case, there are just three alternatives to you being an animal: (1) there is no animal where you are; (2) there is an animal there, but it doesn't think in the way that you do; or (3) there is an animal there, and it thinks exactly as you do, but you are not it. (...) The repugnancy of these alternatives seems to me powerful reason to suppose that you are an animal. » E. Olson (2007), p.29-30.

⁸⁰ E. Olson (2007) p.139-141.

⁸¹ E. Olson (2007) p. 223-232.

de l'organisme biologique sont bien réelles et que ce sont les additions des philosophes dans la tradition de J. Locke et de D. Hume qui ne le sont pas⁸².

T. Nagel affirme que le cerveau est une composante essentielle du moi en s'appuyant sur une théorie causale de la référence. Il relie cette question à la présence d'un « effet que cela fait d'être un moi ». L'âme intervient comme une hypothèse pour rendre compte de ce phénomène. Or, le cerveau peut être également un bon candidat pour y arriver en tant que chose munie de deux propriétés, une physique et l'autre mentale⁸³. D'autre part, nos réactions à la présentation de différents scénarios imaginaires et les méthodes centrées sur la conscience introspective du moi sont à analyser avec prudence, car elles négligent la possibilité d'identifier le mauvais objet⁸⁴. Le mot « je » pointe en direction du monde une réalité que des observations empiriques doivent compléter. Il indique ainsi la cause de « l'effet que cela fait ». L'observation nous révèle que c'est le cerveau⁸⁵. L'impression de pouvoir changer de corps s'explique par *l'ignorance a priori* que nous avons de cette cause, nous permettant d'imaginer plusieurs possibilités. Cependant, lorsque le référent réel est établi, il ne devient plus possible d'imaginer ces situations sans commettre une erreur de référence. Dans ce cadre, désigner une réplique de moi-même comme étant moi constitue un échec et non un choix de description⁸⁶.

E. Olson, et T. Nagel établissent la substantialité du moi en partant de l'observation du monde extérieur et non de la conscience du moi à l'œuvre dans l'imagination. P.

⁸² Q. Cassam (1993) adresse une objection similaire à Parfit en affirmant que sous l'animalisme les frontières du moi sont indépendantes de notre esprit et dépendent d'un fait à propos du monde, contrairement aux autres conceptions, qui sont des produits de notre esprit.

⁸³ T. Nagel (1986), p.28-32.

⁸⁴ T. Nagel (1986), p.32-37.

⁸⁵ « My concept of myself contains the blank space for such objective completion, but does not fill it in. I am whatever persisting individual in the objective order underlies the subjective continuities of that mental lives that I call mine. But a type of objective identity can settle the questions about identity of the self only if the thing in question is both the bearer of mental states and the causes of their continuity when there is continuity. If my brain meets these conditions then the core of the self – what is essential to my existence – is my functioning brain. » T. Nagel (1986), p.40.

⁸⁶ T. Nagel (1986) p.37-45.

Unger reprend quant à lui cette méthode introspective pour en conclure que nous sommes une capacité mentale de base réalisée dans la structure d'une chose physique⁸⁷. Il propose, dans la lignée de B. Williams, d'employer un test nous demandant de choisir entre une douleur présente ou une douleur après une opération imaginaire menaçant notre identité afin de cerner l'objet de nos attitudes prudentielles, par opposition à des attitudes compatissantes, sympathiques ou morales⁸⁸. Selon lui, nous préférerions qu'une copie de nous vive la douleur, de même que nous éviterions une souffrance intense à un être qui résulte d'une perte complète nos états mentaux tant que notre cerveau continue de réaliser un potentiel d'activités mentales de base⁸⁹.

Il suggère que c'est l'interruption du potentiel d'activité mentale dans le processus physique de réplication qui nous amène à ne pas faire de l'être résultant un objet de nos attitudes prudentielles. En effet, lorsque les informations sur l'état de mon corps sont enregistrées dans une machine et transmises par ondes radios sur Mars, ni la machine, ni les ondes radios ne constituent des objets physiques dont la structure permet une activité mentale de base. Il y a donc une interruption physique de cette capacité et elle coïncide avec la disparition de nos attitudes prudentielles⁹⁰.

Comme T. Nagel, B. Dainton définit le moi comme un être pour qui il y a un « effet que cela fait d'être ». Comme R. Swinburne, il estime que sa présence est indiquée dans l'unité phénoménale des expériences en un instant et dans leur brève succession. L'incapacité d'imaginer mon existence aller dans une direction différente de ce flux l'illustre. Comme T. Nagel, il considère le cerveau comme étant le référent de ce sujet d'expérience en tant que chose « phénoménale » et non seulement physique. Afin de rendre compte de notre survie en dépit de l'interruption de ce flux de

⁸⁷ P. Unger (1990), p.102-169.

⁸⁸ P.Unger (1990), p.27-34.

⁸⁹ P. Unger (1990), Chapitre 5.

⁹⁰ P. Unger (1990), Chapitre 4 et 5.

conscience lors du sommeil, il propose de considérer le sujet de ce flux de conscience comme la chose munie du potentiel d'avoir ce genre d'expériences : le cerveau. Dans le sommeil, il le possède de façon ininterrompue⁹¹. Comme P. Unger, il reproche à la réplique de contenir une interruption de ce potentiel. Ni la machine, ni les ondes radios se dirigeant vers Mars ne pourraient produire un « effet que cela fait » d'être une machine ou des ondes radios. Donc, la réplique serait substantiellement un autre sujet d'expérience⁹².

P. Unger et B. Dainton ont en commun de considérer la causalité cérébrale de la vie mentale et les frontières physiques qui la constituent comme étant suffisantes à constituer une chose tangible, concrète, réelle qui agit comme un être ayant un point de vue sur le monde qui l'entoure. Comme E. Olson et T. Nagel, ils s'opposent à D. Parfit en niant que le moi puisse être une diversité de corps reliés ensemble par une psychologie. Le moi est essentiellement une chose et non un ensemble d'événements mentaux inter-reliés.

Qu'en est-il des arguments de Parfit, qui présente en imagination des cas d'indétermination ? P. Unger et Q. Cassam répondent que la possible indétermination d'une chose physique n'en fait pas pour autant une entité conceptuelle produite par notre esprit. Des objets physiques courants ont des frontières mal définies. Leur imprécision n'est pas due à nos concepts, mais à la réalité. Ainsi, même s'il y a indétermination quelque part dans les spectres ou lors de ma division, le moi a des frontières physiques concrètes qui ne s'étendraient pas jusqu'à une copie physique parfaite de mon corps actuel⁹³.

⁹¹ B. Dainton (2004)

⁹² B. Dainton (2008), p.178-200.

⁹³ Q. Cassam (1993), section 4. P. Unger (1990).

1.4- Présence d'une chose dans le flux de conscience

D. Parfit suggère que définir le moi comme une chose physique ou comme un ensemble d'événements mentaux reliés ensemble par des souvenirs, des intentions, des traits de caractère et des croyances qui s'enchaînent graduellement ne repose pas sur un fait. Selon lui, de nombreux auteurs rejettent cette thèse en vertu d'une croyance non réductionniste implicite. Dans la section précédente, nous avons passé en revue différents arguments d'auteurs qui refusent ces affirmations. Ils ne défendent pas que le moi est purement physique, car il lui reconnaisse la présence d'un « effet que cela fait » et lui attribue la pensée, mais ils nient que cette activité mentale puisse être supportée par un sujet pouvant logiquement changer de corps ou de cerveau. De plus, ils affirment que nous survivrions à une altération massive de nos processus mentaux tant que cette chose physique continue de produire une activité consciente.

Leurs arguments sont de deux types. D'une part, il y a ceux qui évitent l'appel à l'introspection en invoquant des considérations sur le monde extérieur. Le corps est l'entité la plus simple à laquelle attribuer des états mentaux afin d'éviter la multiplication des entités superflues ou des erreurs de catégories, tout en respectant l'intuition que notre réalité ne dépend pas d'une vue de l'esprit. Une autre suggestion, c'est que le mot « je » réfère à la cause de nos pensées et les observations nous révèlent que c'est le cerveau. D'autre part, certains suivent B. Williams en récupérant les diverses fictions entourant l'identité du moi. Ils pointent la présence d'une intuition qui correspond avec l'interruption du potentiel de conscience.

Dans ce qui suit, je veux montrer que ces objections échouent parce qu'elles ne tiennent pas compte de l'intuition très forte selon laquelle le moi est une chose *présente à la conscience*. Ni le corps, ni le cerveau ne répondent à cette intuition. C'est pourquoi l'approche centrée sur la poursuite d'événements mentaux représente une alternative sérieuse. Ce qui explique la présence d'une indétermination, c'est que

les deux théories contiennent chacune séparément une partie des intuitions du sens commun, mais que seulement une croyance non réductionniste répond entièrement à ces intuitions. Une théorie de l'identité du moi à travers le temps ne peut donc pas s'appuyer sur l'intuition commune, mais doit en proposer une révision. C'est dans ce cadre qu'il faut comprendre le raisonnement de D. Parfit.

1.4.1- Le corps

E. Olson avance qu'une conception du moi définie par un ensemble de processus mentaux permet de faire cohabiter plusieurs moi potentiels à une seule place. Il reproche à cette hypothèse d'être bizarre. Le corps, en lui-même, ne lui apparaît pas contenir de telles implications. C'est la même bizarrerie qui amenait B. Williams à objecter la possibilité de la duplication. Or, ils n'évitent pas ce problème.

E. Olson a tort de considérer son approche comme étant celle du sens commun. Nous le constatons en relisant le contexte historique décrit précédemment. Autant R. Descartes, J. Locke, D. Hume, B. Williams que leurs héritiers manifestent une attitude commune : ils cherchent la nature du moi au moyen d'exercices d'introspections. Puisque les patterns des philosophes constituent une donnée nous indiquant la nature du sens commun, ce détail a son importance. Qu'est-ce que cela nous révèle ? Qu'il y a une intuition forte selon quoi le moi se caractérise par *la manière dont il reconnaît sa propre existence*. Il apprend son existence par une voie interne et non externe, d'où la tendance marquée d'employer l'imagerie mentale⁹⁴.

De plus, les non-philosophes sont prompts à admettre la possibilité de la réincarnation, de la résurrection et d'une vie incorporelle. Ils se racontent des récits d'esprits qui les hantent. Ils parlent facilement de leur être comme quelque chose qui occupe et non qui est le corps. Le trait commun de ces croyances, c'est que le moi se reconnaît comme étant mental et non physique. Ils ne perçoivent aucune bizarrerie

⁹⁴ D. Dennett (1991) élabore abondamment autour de cette idée.

dans une vie qui se succède dans plusieurs corps. E. Olson est marginal de ce point de vue. La réflexion philosophique peut expliquer ces intuitions. C'est parce que nous semblons reconnaître notre existence par une voie interne que nous nous identifions spontanément comme étant une chose mentale.

Le problème apparaît lorsque nous apprenons que la cause de nos pensées est physique suite à une immersion dans une culture moderne valorisant la science. Cette conversion, en elle-même, stimule souvent un sentiment d'inconfort. C'est pour cela qu'il y a autant d'articles écrits sur le sujet⁹⁵. Apprendre que nous sommes un corps, si nous le sommes, est donc tout autant « bizarre » qu'E. Olson pense qu'il est « bizarre » d'être défini par un ensemble de processus mentaux.

Les diverses approches philosophiques discutées dans ce texte sont des suggestions visant à sortir de cette zone d'inconfort. Dans ce contexte, il y a une raison pour laquelle le moi est plutôt défini par un ensemble de processus mentaux : c'est la meilleure manière de rendre compte de l'intuition selon laquelle le moi reconnaît son existence de l'intérieur, par une voie privée, tout en respectant par ailleurs ce que nous connaissons des causes de ces processus. En comparaison, le corps est une entité dont nous reconnaissons la réalité via des manuels de biologie.

Cela n'empêche pas les raisons d'E. Olson invoquées en faveur du corps d'avoir un certain poids. C'est justement parce qu'il y a des intuitions allant dans les deux directions qu'il y a indétermination.

1.4.2- La cause de nos processus mentaux

L'approche de T. Nagel peut faire l'objet d'une critique similaire. Il fait plusieurs suppositions quant à l'intention que nous visons lorsque nous faisons usage du « je ».

⁹⁵ Voir D. Fiset et P. Poirier (2000) pour une revue des problèmes philosophiques émergents d'une confrontation entre la psychologie du sens commun et le modèle d'explication scientifique de l'esprit humain.

Nous visons la cause de nos pensées et de nos expériences, nous laissons ouverte la possibilité d'identifier la mauvaise personne, nous pointons en direction du monde extérieur. Partant de ces présupposés, il suppose que nos découvertes à propos du cerveau satisfont nos attentes et que, par la suite, nos imageries de changement de corps deviennent des bizarreries vides de sens.

Or, il ne va pas de soi que c'est ce que nous visons lorsque nous usons du « je ». Nous semblons plutôt viser à l'intérieur du flux de conscience et nous fermons la possibilité de nous tromper de cible. Une vaste communauté de penseurs s'oriente vers une réalité intérieure, l'imagerie mentale ou l'auto-observation, plutôt que vers le monde extérieur pour trouver la véritable nature du moi.

D'autre part, lorsque les individus ordinaires font référence à eux-mêmes dans le passé via la mémoire, ils n'étudient pas le phénomène physique ayant mené à cette représentation pour en conclure qu'elle est fidèle et qu'elle est arrivée à eux. Ils rapportent l'information comme s'il voyait directement la scène se passer *dans* leur tête. L'hypothèse qu'elle est arrivée à une autre personne et qu'ils commettent une erreur d'identification leur apparaîtrait étrange. Je soupçonne que c'est parce qu'ils croient voir directement leur présence dans le passé. En ce sens, la possibilité d'une erreur de référence leur semble exclue.

Je formule ici une hypothèse qui reste à vérifier. À défaut d'effectuer une enquête de perception, la méthode cartésienne demeure utile à cet effet. Selon Descartes, je suis assuré de mon existence en analysant mon activité consciente. Dans cet ordre d'idée, puis-je être assuré de mon existence hier ? Se peut-il qu'un scientifique fou ait créé une copie parfaite d'un autre être que moi pendant la nuit et que je sois cette copie ? Personnellement, lorsque je me pose la question, j'ai le réflexe de répondre spontanément non. Pourquoi ? Parce que j'ai l'impression de me voir directement dans le passé. Je n'ai pas l'impression de voir directement mon cerveau dans le passé, mais de me voir moi. En ce sens, en employant le « je », je vise spontanément

une chose à l'intérieur du flux de conscience et non à l'extérieur et je ne reconnais pas la possibilité de me tromper sur la cible de ma référence.

Ces remarques se retournent contre l'objection de T. Nagel, car, lorsque nous découvrons que le cerveau est la cause de tous nos processus mentaux, nous pouvons être insatisfaits de la réponse et trouver qu'elle ne correspond pas tout à fait à la cible visée dans notre acte de référence. Le fait que le cerveau est à l'extérieur de notre flux de conscience peut nous amener à refuser d'y voir ce que nous y cherchions. La duplication pendant la nuit peut illustrer ce point. Si « je » renvoie au cerveau, alors, oui, rien n'empêche d'imaginer que mon cerveau est la copie d'un autre. Pourquoi ? Parce que je ne vois pas, de l'intérieur, mon cerveau et donc je ne peux pas prétendre voir directement s'il était là hier. Je peux simplement conjecturer à partir de mes observations présentes que le monde qui m'entoure a l'air assez normal pour ne pas prendre au sérieux cette hypothèse. Si je dis « je » en visant une chose dont j'ai *directement* conscience, je ne vise donc pas le cerveau.

Ce que je vise peut être une fiction, c'est-à-dire être la chose de plus dont D. Parfit conteste l'existence. Je pointe alors dans le vide. Si c'est le cas, j'ai un choix. Je peux réviser mes attentes relativement au référent du « je » de la manière dont propose T. Nagel et me satisfaire du cerveau. Sinon, je peux ne pas viser « la cause de mes processus mentaux », mais plutôt les composantes qui se rapprochent le plus de la dimension interne de la conscience que j'ai de moi-même, ainsi que du sentiment de certitude qui l'accompagne. De ce point de vue, la conception du moi centrée sur la poursuite d'un ensemble d'événements mentaux constitue une cible de référence beaucoup plus près de mes attentes⁹⁶. Elle vise un phénomène interne et je suis beaucoup plus certain de son occurrence que je le suis de la réalité physique qui

⁹⁶ D. Parfit (1984), p.468-477 répond ainsi à T. Nagel en posant le problème en termes d'un acte de référence partiellement manqué. Face à cette situation, nous avons plusieurs choix quant à la manière de corriger l'intention visée par la référence. L'indétermination entre les deux approches en découle. L'analyse que je fais de l'échec partiel de la référence diffère cependant de celle de D. Parfit.

la supporte⁹⁷. J'ai une relation plus intime et directe à l'apparence de mes souvenirs et de mes intentions qu'à ma neurologie⁹⁸.

1.4.3- L'interruption du potentiel de conscience

E. Olson et T. Nagel commettent l'erreur de sous-estimer la dimension contre-intuitive de leur position en mettant davantage l'accent sur celle de leurs adversaires. P. Unger et B. Dainton font la même chose en prétendant que nous correspondons à une chose dont les frontières sont délimitées par un potentiel ininterrompu de conscience.

Premièrement, ils montrent que leur intuition coïncide avec cette frontière et non qu'elle porte sur elle. Ils visent peut-être la chose de plus, non-réductible, que Parfit les soupçonne de viser. Ils se réfèrent à l'âme en disant se référer au cerveau parce qu'ils croient inconsciemment que la destruction du cerveau causerait la destruction de l'âme⁹⁹. Comment prêter une telle référence inconsciente à des auteurs ? En observant leur comportement. Ils disent pointer un potentiel ininterrompu de conscience, mais ils nous demandent de regarder vers une réalité dans nos pensées et sensations. Ils prétendent indiquer une réalité extérieure, mais nous montre une réalité intérieure. Ils imaginent des scientifiques fous tortionnaires et nous racontent l'effet que cela leur fait comme Descartes se servait du malin-génie pour y découvrir une chose pensante incorporelle. Bref, ils réfèrent à la même chose, mais le nient.

⁹⁷ Ce qui sera l'objet du prochain chapitre.

⁹⁸ Même si cette relation n'est pas directe, car je peux avoir une illusion de souvenir. Cependant, nous pouvons affirmer que j'ai une relation *plus* directe à cet état mental qu'à sa cause physique.

⁹⁹ D. Parfit répond ainsi : « I suspect that reviewing my argument would never wholly remove my doubts. At a reflexive or intellectual level, I would remain convinced that the Reductionist View is true. But at some level I would still be inclined to believe that there must always be a real difference between some future person's being me, and his being someone else. (...) It may help to add these remarks. On the Reductionist View, my continued existence just involves physical and psychological continuity. On the Non-Reductionist View, it involves a further fact. It is natural to believe in this further fact, and to believe that, compared with the continuities, it is a *deep* fact, and is the fact that really matters. *When I fear that, in Teletransportation, I shall not get to Mars, my fear is that the abnormal cause may fail to produce this further fact. As I have argued, there is no such fact. What I fear will not happen, never happens.* [je souligne] » D. Parfit (1984), p.279.

Deuxièmement, nous pouvons raffiner l'exercice d'introspection de manière à montrer comment c'est le cas. Si nous apprenions de la part de physiciens qu'à l'occasion, suite à un phénomène étrange, notre cerveau se transforme pendant la nuit en une matière complètement inapte à produire la moindre activité consciente (aussi inapte que peut l'être une machine à répliquer ou un pattern d'ondes sonores dans l'air), comment réagirions-nous à cette nouvelle ? Je doute que nous en tirerions la conclusion qu'un autre être parfaitement similaire à nous est mort pendant la nuit, puis que nous avons pris sa place. Pourtant, c'est précisément ce que leur théorie implique. Cela ne correspond pas à nos intuitions. Notre intuition est que la référence du moi ne se trouve pas dans le monde extérieur, mais à l'intérieur du flux de conscience. Or, l'interruption du potentiel de conscience est un fait extérieur et c'est pourquoi il n'a pas d'importance dans ce scénario.

Cette idée peut être clarifiée en combinant l'imaginaire du zombie à celui de la réplication. Le zombie renvoie dans la philosophie anglo-saxonne à la possibilité qu'un être puisse manifester tous les signes extérieurs de la conscience sans en avoir une. Par exemple, il hurle « j'ai mal » en se tordant, mais il n'y a aucun effet que cela fait pour lui d'être dans cette situation. Dans ce cas, l'effet de la douleur est une entité dont la réalité ne peut être pointée qu'intérieurement par des êtres qui en font l'expérience¹⁰⁰. Suivant la même logique, la destruction de mon corps suivie de la création d'une copie parfaite de moi peut être analysée comme un cas de zombie partiel. La succession de corps manifestera les signes extérieurs d'un moi identique à travers le temps, mais, en réalité, il ne fera pas l'*expérience* de cette identité. La supposition sous-jacente, c'est que, contrairement à cette entité, nous en faisons l'expérience. L'imaginaire de la réplication, comme celle du zombie, intervient comme moyen de pointer cette réalité *intérieure* qui nous constitue. C'est peut-être cela que P. Unger et B. Dainton cherchent. Si c'est le cas, cette thèse est non réductionniste. Le réductionniste, lui, affirme que nous sommes, *intérieurement*,

¹⁰⁰ S. Blackburn (1999), p.71-77.

exactement comme les répliques. La seule chose qui nous distingue, ce sont des propriétés extérieures (la discontinuité physique dans les causes de cette intériorité). Or, ce n'est intuitivement pas cela que nous visons par « je ».

D. Parfit reconnaît que la réplication contient une dimension contre-intuitive, mais il l'explique par la présence d'une illusion. B. Dainton et P. Unger y succombent peut-être sans s'en rendre compte.

1.4.4- Révisionnisme ontologique

Dans ce qui précède, nous avons précisé ce en quoi consiste le réductionnisme de D. Parfit quant à l'identité et l'existence du moi, ainsi que ses arguments. Par la suite, nous avons vu des objections visant à nier que le choix entre une conception du moi comme chose physique singulière, que ce soit le corps ou le cerveau et une conception centrée sur la poursuite d'états mentaux comme la mémoire, les intentions et autres est indéterminée. Il n'y aurait aucun choix. Nous sommes une chose singulière et l'approche de la continuité psychologique est fausse en vertu de la réalité. Par la suite, j'ai montré comment ces objections négligent des intuitions fortes selon lesquelles le moi est une chose *présente* à la conscience. Je conclurai en établissant le lien avec la réflexion de D. Parfit.

Sa pensée commence avec une affirmation apparemment banale. Le moi consiste en l'existence d'un corps, d'un cerveau et d'une série d'événements physiques et mentaux inter reliés et non une chose de plus. Présenté ainsi, il est difficile de savoir où nos intuitions se situent, car l'affirmation demeure assez abstraite.

Il avance un test pour vérifier si nous sommes d'accord avec cette affirmation : la possible indétermination de l'identité du moi à travers le temps. Il imagine des situations où un composant accompagnant normalement le moi est soustrait, alors que les autres sont présents. Le cerveau est transplanté dans un autre corps. Le corps est

détruit et remplacé par une copie similaire qui a des états mentaux indiscernables des nôtres. Le flux de conscience est divisé en deux par double transplantation d'hémisphères cérébraux équipotents. Le corps est reconstruit partiellement avec des cellules nouvelles. Face à ces éventualités, le réductionniste doit logiquement réagir en y voyant un faux débat à propos de la manière d'étendre un concept flou à une situation pour laquelle il n'a pas été prévu. Pourquoi ? Parce qu'il a devant lui tous les composants du puzzle. À l'inverse, celui qui est insatisfait de cette réponse et qui continue d'en chercher une en multipliant les arguments et les exercices d'introspection trahit son adhésion implicite au non réductionnisme, car il estime alors qu'il n'a pas toutes les composantes du puzzle devant lui et qu'il lui manque l'ingrédient central : le moi.

Ce que nous venons de voir, c'est qu'il y a des penseurs qui nient être des non réductionnistes, mais qui réagissent à l'éventualité de ces scénarios de la manière dont Parfit estime qu'elle révèle une adhésion implicite au non réductionnisme. Ils nient principalement que le moi puisse renvoyer à un ensemble d'événements mentaux inter-reliés plutôt qu'à une chose physique singulière. Ensuite, ils emploient des exercices d'introspection ou fouillent dans le monde extérieur pour nous le montrer. Succombent-ils à une telle illusion ?

La raison principale qui amène Parfit à douter que le « je » renvoie à une chose physique en particulier, c'est que nos attitudes relativement à la présence du moi dépendent de « l'effet » et non de la cause. « L'effet », c'est la continuité d'un processus mental. Or, le corps n'en est que la cause et donc nous pourrions continuer d'employer le « je » si ce processus était plutôt poursuivi par une machine à répliquer. C'est pourquoi les événements mentaux inter-reliés sont concurrents au corps ou au cerveau comme référent du « je » et qu'il y a indétermination¹⁰¹.

¹⁰¹ D. Parfit (1984), p.286.

Ce qui rend le problème confus, c'est que nous ne référerions probablement pas à la copie de nous en disant « je ». Les objections précédentes en témoignent. Selon Parfit, c'est parce qu'aucune forme de réductionnisme, qu'il soit centré sur une chose physique ou sur un ensemble d'événements mentaux, ne répond à nos intuitions. Celles-ci sont non réductionnistes. J'ai clarifié ce point en montrant comment nous visons une chose présente dans le flux de conscience en employant « je ». À cet égard, le corps ou le cerveau sont des référents tout autant contre-intuitifs.

Rendu à ce stade de la réflexion, il est possible de clarifier en quoi les deux versions du réductionnisme sont contre-intuitives. Se référer à un ensemble d'événements mentaux *uniques* pour renvoyer à nous implique aussi de pointer une réalité à l'extérieur du flux de conscience. C'est ce que l'argument du transplant de trace mnésique, le cogito révisé, le spectre combiné, l'examen de math et la fission illustrent. Je ne peux pas voir si le passé est arrivé à moi via la mémoire, car mon état mental m'a peut-être été transplanté, je suis peut-être une copie de la personne qui l'a vécu ou je suis peut-être le résultat d'une double transplantation d'hémisphères cérébraux. Je ne peux pas exclure ces hypothèses en observant à l'intérieur de mon flux de conscience. Je dois regarder le monde extérieur et constater l'absence de telles bizarreries. Ce qui est étrange dans la fission, selon les auteurs qui en discutent, c'est précisément que je puisse disparaître alors que le flux de conscience qui me constitue, lui, continue comme à l'habitude.

Lorsque nous reconnaissons qu'aucun réductionnisme ne satisfait cette intuition, nous devons en conclure que le « je » courant contient une part d'illusion. Face à ce constat, nous sommes contraints d'abandonner une partie de notre intuition. C'est alors que les approches centrées sur une chose physique singulière et sur un ensemble d'événements mentaux sont des candidats potentiels et indéterminés.

Cela étant dit, il y a une raison de privilégier plutôt l'approche centrée sur la continuité psychologique : c'est celle qui se rapproche le plus de l'idée d'une

présence à soi dans le flux de conscience. Nous avons une relation plus directe, plus intime, plus spontanée, plus certaine à nos processus mentaux qu'à notre corps.

CHAPITRE II

AUTORITÉ PRIVILÉGIÉE DE SOI SUR SON PROPRE PASSÉ

Au chapitre précédent, nous avons clarifié la conception réductionniste du moi de D. Parfit, qui pose que notre existence consiste en un corps et un ensemble d'événements physiques et mentaux inter-reliés et non en une chose de plus. Selon lui, l'opposition entre une conception centrée sur une chose physique ou sur la poursuite d'une psychologie est indéterminée. Notre inconfort à le reconnaître trahit la croyance en une chose de plus que les entités reconnues par le réductionnisme. J'ai affirmé, en réponse à des objections substantialistes, qu'il y a une telle part d'illusion. Elle est l'intuition d'une présence de soi dans la conscience. Aucune des deux approches n'y répond. De cet échec découle l'indétermination.

Dans ce qui suit, nous clarifierons un autre outil que D. Parfit déploie en faveur du réductionnisme : la thèse d'une description adéquate de la vie mentale qui omet la référence à soi. Nous expliquerons le problème de circularité qu'elle vise à surmonter, en analysant les diverses fictions qu'il invoque pour rendre intelligible ce point de vue. Ensuite, nous opposerons cette thèse à une objection. Selon certains, la conscience de soi ne se caractérise pas par sa présence, mais par son absence du processus de la pensée. Nous ne reconnaissons pas intérieurement un moi avant de conclure à son identité à travers le temps. Le contexte normal nous dispense d'avoir à le faire. C'est ce qui définit une perspective à la première personne. Or, la démarche de D. Parfit déforme cette perspective. Je montrerai comment cet aspect se ramène à la spontanéité de la représentation et ne définit pas la perspective du moi, contrairement à une illusion non réductionniste.

2.1- Dépersonnalisation de la mémoire

2.1.1- Problème de circularité

La compréhension de la mémoire nécessite de se référer à un seul et même moi au fil du temps. En effet, pour reconnaître la représentation d'une expérience passée comme étant un souvenir plutôt qu'une autre forme de connaissance ou une illusion, l'expérience doit être arrivée à moi et non à un autre. Autrement dit, la conscience unifiant ma vie est un point de vue à partir duquel *un même être*, moi, se représente lui-même comme étant lui-même à plusieurs instants¹⁰². C'est en ce sens que J. Locke définissait l'identité personnelle par la conscience comprise comme faculté de se « (...) consulter soi-même comme le même, comme une même chose qui pense en différents temps et en différents lieux ». Cette dépendance conceptuelle pose un problème de circularité à ceux qui veulent, dans la tradition de J. Locke, se servir de la mémoire pour analyser ce en quoi mon existence consiste. L'élément de l'analyse (la mémoire) contient le terme à analyser (le même être à travers le temps : moi)¹⁰³.

2.1.2- Thèse de la description complète impersonnelle

Afin de surmonter cette difficulté, D. Parfit imagine un point de vue quasiment identique à celui impliqué dans la mémoire, mais qui ne nécessite pas la présence d'un même être pour être compris. Autrement dit, il construit une perspective sur l'unité consciente de la vie quasiment à la première personne, mais qui ne stipule pas la présence d'un même moi. Le passage d'un état mental à un autre a la même apparence et la même cause qu'en temps normal, mais rien n'indique si un, deux ou aucun être (s) ne se le représente (nt). Il qualifie ce processus de « quasi-

¹⁰² J. Locke (2009), p.522.

¹⁰³ Cette difficulté prend sa source de J. Butler, qui, s'adressant à J. Locke, affirme « And one should really think it Self-evident, that Consciousness of personal Identity presupposes, and therefore cannot constitute, personal Identity, any more than Knowledge in any other case, can constitute Truth, which it presupposes. » J. Butler (1736), p.302.

mémoire »¹⁰⁴. Il n'explicite pas les caractéristiques de cette apparence de la mémoire, mais ses exemples suggèrent une représentation sensorielle du passé centrée autour d'une position dans l'espace, fidèle, autoréférentielle, riche en détail et vive. Sa cause est une trace physique laissée dans le cerveau par l'expérience¹⁰⁵. Cette relation ne stipule pas si le sujet de la représentation et le sujet de l'expérience sont le même. D. Parfit va jusqu'à imaginer qu'une quasi-mémorisation puisse se produire en faisant abstraction complètement de l'existence d'êtres qui se la représentent¹⁰⁶. Il défend l'intelligibilité de cette perspective en la mettant en scène dans des fictions.

2.1.3- Arguments

Premièrement, il imagine que la trace physique laissée par le voyage de Paul à Venise dans son cerveau est copiée et transplantée dans le cerveau de Jane. Celle-ci se représente alors de façon détaillée l'expérience de Paul du point de vue de Paul. Sachant que sa représentation est causée par cette expérience, Jane peut donc savoir l'effet qu'un tel voyage faisait à Paul et donc être reliée à son passé par un processus similaire à la mémoire intérieurement et extérieurement sans pourtant avoir été celle qui l'a vécu¹⁰⁷. L'apparence et la cause de la mémoire peuvent donc se transmettre d'une personne à une autre et ne nécessitent pas de se référer au même moi¹⁰⁸.

¹⁰⁴ « To answer this objection, we can define a wider concept, *quasi-memory*. I have an accurate quasi-memory of a past experience if (1) I seem to remember having an experience, (2) someone did have this experience, and (3) my apparent memory is causally dependent, in the right kind of way, on that past experience. », D. Parfit (1984), p.220.

¹⁰⁵ D. Parfit (1984), p.220-222.

¹⁰⁶ D. Parfit (1984), p.225-226 et D. Parfit (1999).

¹⁰⁷ D. Parfit (1984), p.220-222.

¹⁰⁸ « (...) quasi-memories would provide a similar kind of knowledge about people's past lives. They would provide knowledge of what these lives were like, *from the inside*. (...) Because we do not have quasi-memories of other people's past experiences, our apparent memories do not merely come to us in the first-person mode. They come with a belief that, unless they are delusions, they are about our own experiences. But, in the case of experiences-memories this is a separable belief. If like Jane we had quasi-memories of other people's past experiences, these apparent memories would cease to be automatically combine with this belief. » D. Parfit (1984), p.222.

Deuxièmement, il invoque l'histoire d'une fission. Notre flux de conscience et le reste de notre psychologie se poursuivent parfaitement en double suite à la transplantation de chacun de nos hémisphères cérébraux dans deux corps similaires au nôtre. La situation peut être présentée de deux manières. Soit nous vivons alors à deux endroits en même temps, soit nous mourons et sommes remplacés par deux êtres qui se représentent notre vie comme nous grâce à la même cause que nous (le même hémisphère cérébral). Selon D. Parfit, ce sont deux descriptions d'une seule et même réalité et non deux hypothèses différentes relativement à ce qui se passe¹⁰⁹. Ce scénario met en scène une perspective quasi-mémorielle de deux manières. D'abord, nous pouvons imaginer entièrement de quoi le processus aura l'air sans décider si les souvenirs qui auront lieu après l'opération seront les nôtres¹¹⁰. D'autre part, si la description choisie est l'apparition de deux nouveaux êtres, cela illustre comment une représentation similaire à la mémoire peut unir plusieurs sujets et non un même moi à travers le temps¹¹¹.

Troisièmement, D. Parfit imagine un univers alternatif et similaire au nôtre où les pensées et les expériences ne sont pas attribuées à des êtres, mais contenues dans des séquences d'événements mentaux¹¹². Il propose alors de remplacer « je » par « ceci » et « tu » par « cela » afin de construire une perspective sur la vie qui ignore complètement l'existence d'un moi tout en distinguant entre un acte d'autoréférence et d'autres formes de références. Suivant cette logique, l'énoncé « je mémorise telle expérience passée » est remplacé par « dans *ceci*, la mémorisation de cette expérience a lieu ». L'énoncé « tu mémorise telle expérience passée » l'est par « dans *cela*, la

¹⁰⁹ En fait, sa position est que la description impliquant la mort suivi de la création de deux êtres est la plus conforme à l'usage du concept de personne, mais il reconnaît que la description impliquant que je vive alors à deux places en même temps est défendable. D. Parfit (1984), p.256-258.

¹¹⁰ D. Parfit spécifie qu'aucune observation interne et externe ne peut déterminer comment appliquer le concept de « moi » dans la situation et qu'avant même de trancher la question, nous saurions que la continuité psychologique à l'œuvre serait parfaitement similaire à celle qui unit nos vies habituellement. D. Parfit (1984), 258-260.

¹¹¹ D. Parfit (1984), p.260.

¹¹² D. Parfit (1999).

mémorisation de telle expérience a lieu ». « Ceci » et « cela » agissent comme pronoms situant intérieurement ou extérieurement les perspectives relativement à des séquences d'événements. Leur usage omet les êtres qui ont ces perspectives¹¹³. En comparant les deux schèmes, il est plus facile d'y voir deux manières de décrire la même réalité et de comprendre la vie mentale adéquatement sans se référer à soi.

2.1.4- Implications sur le problème de la nature du moi

Par ce détour, D. Parfit contourne un problème de circularité. Puisque qu'une relation unissant divers états mentaux de façon similaire à la mémoire peut être comprise sans se référer à l'identité du moi à travers le temps ou même à son existence, il est alors possible d'analyser le moi comme un ensemble composé de relations unificatrices de ce genre. Nous ne retrouvons donc plus dans l'analyse ce qui est à analyser : le moi. La circularité disparaît¹¹⁴.

Ensuite, il utilise cet exercice comme moyen de clarifier son réductionnisme. Il soutient que l'existence et l'identité du moi à travers le temps « consistent » en un ensemble de faits plus particuliers : l'existence du corps et l'occurrence d'une série d'expériences et de pensées reliées ensemble par la mémoire, la poursuite d'intentions, la persistance de traits de caractère et de croyances. Le verbe « consister » signifie alors que le passage du point de vue dépersonnalisant ces faits à un point de vue qui reconnaît la présence d'un moi n'indique rien de plus qu'une nouvelle description pour en parler et non une réalité qui les transcende¹¹⁵.

¹¹³ « In place of the pronoun « I, » these beings might have a special use of “this” which referred to the sequence in which this use of “this” occurred. Where one of us would say, “I saw the Great Fire,” one of them would say “This included a seeing of the fire”. In place of “you,” they might have a special use of “that”, which refer to the sequence to which it was addressed. Where we would say, “Did you see the Great Fire?” they would say, “Did that include a seeing of the fire?” D. Parfit (1999), p.229.

¹¹⁴ D. Parfit (1984), p.222, 226 et D. Parfit (1999), p.222.

¹¹⁵ D. Parfit (1984), 223, D. Parfit (1995), p.19-20 et D. Parfit (1999), p.261.

Finalement, D. Parfit fait ressortir une croyance non réductionniste non reconnue par de nombreux philosophes traitant de ces questions. Il peut nous apparaître que « quelque chose manque » dans la description alternative de la vie mentale qu'il nous offre. Nombreux sont ceux qui cherchent dans l'expérience consciente, dans l'imagination ou dans la réalité extérieure plutôt que simplement dans le langage cet ingrédient absent. Cela trahit la croyance qu'au-delà du corps et des événements mentaux, il y a une chose inanalysable qui correspond à moi, même si ces auteurs prétendent le contraire¹¹⁶.

2.2- La mémoire comme source d'information fiable et spontanée

Afin d'évaluer la justesse de cette position, il faut déterminer si la perspective quasi mémorielle présentée par D. Parfit ressemble à la mémoire normale sous ses aspects les plus distinctifs, car c'est à cette condition que nous pouvons présenter les deux perspectives comme portant sur une seule réalité et non deux réalités distinctes. Plusieurs auteurs nient cette ressemblance sans pour autant postuler l'existence d'une chose spéciale qui transcende le corps, les expériences et les pensées particulières, ainsi que le langage pour en parler. Leurs critiques posent que la présence d'un moi identique à travers le temps s'exprime dans un contexte où l'information à propos du passé est introduite spontanément, par opposition à des contextes où elle nécessite une recherche active¹¹⁷. D. Parfit élimine cette distinction en traitant le moi comme quelque chose dont il faut établir la présence, soit à l'intérieur du flux de conscience, soit à l'extérieur dans des relations causales. Or, le moi ne se cherche pas. Sa présence se présume.

¹¹⁶ « In our thoughts about our own identity, we are prone to illusions. That is why the so-called "problem cases" seem to raise problems: why we find it hard to believe that, when we know the other facts, it is empty or a merely verbal question whether we shall still exist. Even if we accept a Reductionist view, we may continue, at some level, to think and feel as if that view were not true." Parfit (1995), p.45.

¹¹⁷ T. Burge (2003), G. Evan (1982), A. Hamilton (2010), Q. Cassam (1999), J. McDowell (1997) formulent des critiques à l'endroit de Parfit qui vont en ce sens.

2.2.1- Immunité à l'erreur d'identification

G. Evan et A. Hamilton attribuent à la mémoire une immunité à l'erreur d'identification¹¹⁸. Ils ne défendent pas, comme Descartes, l'accès à notre existence passée comme étant certaine, mais affirment plutôt que, pour avoir la forme d'une connaissance des expériences passées comparable à la mémoire, une représentation ne peut se présenter comme étant véridique en laissant ouverte la question de l'identité des sujets concernés¹¹⁹. Dans une perspective à la troisième personne, je peux me représenter une silhouette au loin comme étant la preuve qu'une personne se trouve à plusieurs centaines de mètres de moi, puis me demander ensuite qui est cette personne. J'ai donc accès à une représentation véridique que je peux ensuite compléter d'une information sur les identités concernées. Dans ce cas, je peux croire qu'une personne est là, mais me tromper sur qui elle est. Ce n'est pas la silhouette de ma copine, mais celle d'une étrangère. Je commets alors une erreur d'identification qui contient malgré tout une part de connaissance¹²⁰. Dans une perspective à la première personne, les informations ne se présentent pas de cette manière. Je n'observe pas la trace d'une expérience dans mon cerveau en produire une représentation fidèle, pour me demander à qui elle est arrivée, puis en conclure que c'était à moi. Comme je n'effectue pas ce mouvement de réflexion, je ne peux donc pas me tromper sur l'identité de celui qui est représenté dans ma mémoire¹²¹.

¹¹⁸ G. Evan (1982) et A. Hamilton (2010). La réflexion sous-jacente s'appuie des distinctions avancées par S. Shoemaker (1970).

¹¹⁹ « L'immunité à l'erreur de mauvaise identification s'exemplifie dans les jugements mnésiques personnels, tels que « J'ai vendu des valises Samsonite au Barkers de Kensington ». Si le locuteur admet qu'il exprime un jugement mnésique personnel, alors bien qu'il est concevable qu'il soit faux, cela ne fait aucun sens pour lui de supposer que ce soit à cause d'une erreur à propos de qui a vendu des valises Samsonite. Si, pour une quelconque raison, j'en viens à douter que j'ai eu ce travail dans ce grand magasin, cela ne fera aucun sens pour moi de continuer à soutenir, en utilisant la même justification, qu'il y a eu cependant quelqu'un qui, jadis, vendait des valises Samsonite. », Hamilton (2010), p.105-106.

¹²⁰ L'exemple m'est inspiré de J. Pryor (1999), p.297.

¹²¹ C'est dans cet ordre d'idée que J. McDowell affirme « Contrast keeping one's thought focused on an ordinary object of perception over a period; this requires a skill, the ability to keep track of

Cette objection vise D. Parfit, car ses arguments ont la forme d'une possible erreur d'identification. Je dois m'imaginer un être qui regarde mon passé exactement de la même manière que moi, qui en a donc une connaissance fiable, mais qui doute seulement de la pertinence d'appeler ce passé le mien ou celui d'un autre, de peur de commettre une erreur d'identification. C'est le cas des deux êtres après ma duplication.

2.2.2- Illusion ou hypothèse confirmée de l'extérieur

J. McDowell reproche ainsi à D. Parfit de faire intervenir comme arguments des scénarios où la connaissance des expériences passées a la forme d'une hypothèse ou d'une illusion. Lorsque Jane a une représentation d'un voyage à Venise, elle déduit d'une source extérieure (tous les détails relatifs à une transplantation de traces dans son cerveau) que le voyage a eu lieu de cette manière. Elle peut croire avoir été à Venise et recevoir ensuite une rétroaction l'informant qu'elle a commis une erreur d'identification, que le voyage a bien eu lieu, mais que c'est Paul qui y est allé. Dans les deux cas, l'état mental de Jane ne ressemble pas à une mémoire, car il ne contient pas l'intériorité propre à la mémoire, qui consiste à se représenter la scène spontanément et fidèlement sans passer par le monde extérieur¹²². En suivant cette logique, dans ma fission, les êtres doivent savoir qu'une double transplantation d'hémisphères cérébraux a eu lieu pour dissocier le processus de quasi-mémorisation de mon identité à travers le temps et apprendre qu'ils ont commis une erreur d'identification. Sinon, ils auront l'impression fausse d'avoir vécu mes expériences. Ici aussi, l'état mental cesse d'être une connaissance immédiate du passé, mais devient une hypothèse confirmée de l'extérieur ou une illusion et perd donc les caractéristiques centrales de la mémoire.

something (...) Continuity of « consciousness » involves no analogue to this: no keeping track of the persisting self that nevertheless seems figure in its content.", J. McDowell (1997), p.232.

¹²² J. McDowell (1997), p.238-242. A. Hamilton (2010) reprend le même argument, p.122-123.

2.2.3- Ensemble de présomptions acquises permettant l'action

T. Burge définit la perspective du moi dans la mémoire par le rôle qu'elle a dans un contexte. Le moi renvoie au centre à partir duquel les expériences, les actions et les pensées sont coordonnées. De plus, si ce centre est un « moi », c'est parce qu'il a une dimension active. Il impose une direction. Or, pour qu'une telle fonction puisse être remplie, un ensemble de connaissances doit avoir été établi préalablement de manière à être utilisé automatiquement sans avoir constamment à douter de sa véracité. Le moi dans la mémoire exprime donc la capacité d'user d'informations fiables à partir du passé pour orienter l'action¹²³. Dans ce cadre, son identité à travers le temps est une condition permettant de passer automatiquement d'une étape à l'autre d'un raisonnement¹²⁴. Selon lui, les êtres qui apparaissent après ma fission ne peuvent automatiquement conclure que leurs représentations sont fidèles à mon passé, mais doivent enquêter sur la fiabilité des transplantations d'hémisphères cérébraux. Autrement dit, dès que j'apprends la nouvelle selon laquelle mon état mental ne me relie pas à mon passé, mais au passé de quelqu'un d'autre, je ne suis plus autorisé à conclure automatiquement à la fidélité de ma représentation¹²⁵.

2.2.4- Moi comme immersion dans un contexte autorisant un type de connaissance

Ces différents arguments se résument à définir le moi par le contexte à l'intérieur duquel une chose est immergée. Ce contexte rend possible un point de vue sur le passé qui est immunisé contre l'erreur d'identification, qui le représente directement sans s'appuyer sur un extérieur et qui permet de passer de façon assurée d'une étape à l'autre d'un raisonnement centré sur l'action. Bref, il permet d'unifier nos vies en se

¹²³ « Exercise *de se* forms [représentations indexées de façon égocentrique, donc à la première personne] or presumptions necessarily requires an explanatory relevant core of veridical applications of such form of coordinates instantiations of psychological states within the same individual to preserve the individual's needs, aims over time." T. Burge (2003), p.297.

¹²⁴ T. Burge (2003), p.302-304.

¹²⁵ T. Burge (2003), p.316-321.

servant du passé spontanément de façon fiable. Être un moi, c'est être situé dans un tel contexte. L'erreur de D. Parfit, c'est de décrire notre conscience du passé en nous demandant de regarder notre vie face à une situation anormale où l'identité impliquée dans le processus est à établir. Le point de vue à partir duquel il analyse le moi ne lui permet donc pas d'en faire une description fidèle à la réalité¹²⁶.

L'objection évite de se positionner entre le réductionnisme et le non réductionnisme.

D'un côté, le moi ne peut pas être analysé comme l'addition d'un langage sur un ensemble de parties plus fondamentales, car les expériences et les pensées particulières sont autant dépendantes du tout qu'elles constituent pour être telles qu'elles sont que l'inverse. Il n'y a donc pas un des niveaux de description qui vient avant l'autre dans la composition de cette réalité. En ce sens, il n'y a pas de réduction à faire de l'un à l'autre¹²⁷.

D'un autre côté, le moi n'est pas non plus une chose que nous observons intérieurement de façon privilégiée à la manière de Descartes et qui transcende le corps, les pensées, les expériences et le langage. L'apparente certitude de ma propre existence est simplement l'expression de la spontanéité avec laquelle je peux en rapporter des informations. Or, c'est un contexte objectif à l'extérieur de ma conscience qui a rendu possible cette caractéristique. L'existence d'un corps humain qui suit une trajectoire stable dans l'espace y contribue fortement. Il sert donc de cible aux actes d'autoréférence¹²⁸. Qu'il puisse disparaître sous l'œuvre d'un malin-génie ou d'un clonage futuriste sans affecter le flux de conscience ne constitue pas une objection au fait que nous sommes ce corps vu d'une certaine manière, car de tels

¹²⁶ "(...) les souvenirs personnels ont une structure holistique : ils ne sont pas une collection d'objets indépendants se rapportant à des événements isolés. Si j'en viens à croire que j'ai été frappé par plus qu'une occurrence limitée de quasi-souvenir, j'en viendrais à me méfier de l'ensemble de mes souvenirs d'enfance dans son intégralité." A. Hamilton (2010), p.122. Voir aussi J. McDowell (1997), p.242-244 et T. Burge (2003), p.326.

¹²⁷ J. McDowell (1997), p.246. T. Burge (2003), p.327.

¹²⁸ C'est la thèse de G. Evan (1982) et de Q. Cassam (1999).

scénarios bizarres rendraient toute notion de moi incompréhensible en éliminant la spontanéité avec laquelle nous pensons.

En fait, J. McDowell reproche à D. Parfit de commettre une erreur similaire à celle de Descartes en cherchant à isoler le flux de conscience du contexte objectif qui l'entoure. Descartes le fait en imaginant qu'un malin-génie puisse le duper sur l'existence de son corps. D. Parfit le fait en imaginant qu'un scientifique « fou » multiplie par deux ses flux de conscience, mélange son être à celui d'un autre, crée des copies parfaites de lui. Autant chez Descartes que chez lui, la pensée dissocie le moi du corps en sortant l'expérience vécue de son contexte corporel normal, puis la met en scène dans un univers où elle en est libérée¹²⁹.

Ce qui distingue D. Parfit de Descartes, c'est qu'il ne défend pas une conception substantialiste du moi par ce détour, mais en fait plutôt une chose superflue. Il dissocie la conscience du corps pour montrer comment autant la conscience que le corps peuvent se passer du « moi » pour être tels qu'ils sont. En imaginant ma conscience se poursuivant dans un autre corps, je vois mon corps sans l'associer à moi¹³⁰. En divisant ma conscience et mes hémisphères cérébraux, je dissocie ma conscience de moi¹³¹. Le moi se révèle donc être une entité optionnelle pour comprendre autant le corps que la conscience¹³².

Les objections mentionnées dans ce chapitre posent qu'en agissant ainsi, Parfit ignore la normalité du contexte, qui nous dispense de douter ainsi de notre réalité. Sa stratégie n' imagine donc rien qui ressemble à notre flux de conscience.

¹²⁹ J. McDowell (1997), p.243-244.

¹³⁰ D. Parfit (1984), p.289-293.

¹³¹ D. Parfit (1984), p.255-261.

¹³² D. Parfit (1984), p.261-266.

2.3- Superfluité du moi dans la mémoire

Afin de faire consister le moi en un corps et/ou en des processus mentaux, D. Parfit surmonte un problème de circularité. Un processus comme la mémoire ne peut apparemment pas analyser le moi, car le fait qu'une expérience soit arrivée à moi est une condition pour que je puisse la mémoriser. Le concept à analyser se retrouve donc parmi les éléments de l'analyse. Parfit répond à ce dilemme que nous pouvons comprendre des états mentaux quasiment identiques à la mémoire, mais qui ne font pas référence à l'identité du moi à travers le temps. Il décrit un point de vue ayant la même apparence et les mêmes causes que la mémoire, mais qui peut se transmettre de moi à un autre. Ce point de vue peut aussi être inclus dans une séquence d'événements mentaux qui contient un acte d'autoréférence effectué à l'aide du pronom « ceci ». Il met en scène des fictions qui en illustrent la possibilité. Cet exercice lui sert à préciser le sens de son réductionnisme, puis à faire ressortir la présence d'une croyance non réductionniste implicite.

Selon certains, cette démarche échoue, car l'état mental que décrit D. Parfit sans référence à l'identité du moi à travers le temps ne ressemble pas à la mémoire. Il lui manque une spontanéité rendue possible par le contexte normal actuel, qui implique la trajectoire d'un objet stable à partir duquel la scène est perçue. Dans la mémoire, l'information ne se présente pas d'abord comme étant véridique pour ensuite exiger une recherche sur l'identité du sujet de l'expérience. En ce sens, l'identité de ce sujet est immunisée contre l'erreur d'identification. Ensuite, le contenu de la mémoire n'a pas la forme d'une hypothèse confirmée par un rapport venant du monde extérieur, mais plutôt celle d'une information à propos du passé qui se présente directement. Finalement, le moi dans la mémoire exprime un contexte où des connaissances ont été suffisamment établies pour ne plus être questionnées. Tous les scénarios avancés par D. Parfit ont en commun de perdre ces caractéristiques. Son erreur est d'analyser le moi en faisant abstraction du contexte à l'intérieur duquel nous en faisons usage.

Cette objection ne repose pas sur la croyance en un fait au-delà du corps, des processus mentaux et du langage pour en parler.

Afin d'évaluer la justesse de ces objections, il faut se demander si un processus aussi spontané et fiable que la mémoire sur les aspects mentionnés plus haut peut se passer de l'identité du moi à travers le temps. Si nous parvenons à le faire, le point de vue impliqué dans ce processus sera quasiment à la première personne et illustrera comment le moi est une entité superflue. Des arguments invoqués contre l'immunité à l'erreur d'identification dévoilent cette superfluité et donc appuient le réductionnisme de D. Parfit¹³³.

2.3.1- Caractéristique non spécifique à la première personne

Premièrement, la spontanéité de la mémoire n'a rien de spécifique à une perspective à la première personne et ne peut donc lui servir de marqueur. Il est vrai que, dans la mémoire, nous n'avons pas d'abord la représentation d'une expérience passée, nous ne cherchons pas activement à qui elle est arrivée et n'en concluons pas finalement que c'est à nous. Cependant, le même genre de spontanéité existe aussi dans une perspective à la troisième personne.

R. J. Howell objecte ainsi à l'immunité à l'erreur d'identification que l'énoncé « ce poisson est un poisson doré et rose » est tout autant immunisé contre elle. Nous ne pouvons pas savoir qu'un poisson est doré et rose et simplement se tromper en pensant que c'est ce poisson-là. Autrement dit, il n'y a pas d'abord une représentation d'un poisson doré et rose suivi d'un questionnement sur lequel il est. Je ne peux me tromper seulement sur son identité en ayant raison sur tout le reste.

¹³³ Je m'appuierai sur les réflexions de J. Pryor (1999), C. Burford (2009), A. Collivia (2006), R. J. Howell (2007) et C. Rovane (1993).

L'immunité à l'œuvre est pourtant la même que celle attribuée à la mémoire dans les critiques précédentes¹³⁴.

Cette réplique est d'autant plus pertinente que Parfit propose justement de substituer le « je » par des pronoms démonstratifs comme « cette », « ceci » et « cela » afin d'éviter de supposer la présence d'un moi. Ainsi, l'énoncé « *Ceci* est une séquence qui contient la mémorisation d'une expérience passée » est immunisé contre l'erreur d'identification exactement de la même manière que l'énoncé « je mémorise une expérience passée ». La mémorisation n'apparaît pas d'abord dans la séquence comme étant fiable pour ensuite être suivi d'un questionnement sur le référent réel de la séquence désignée par « ceci »¹³⁵. Lorsque D. Parfit substitue un pronom par l'autre, il peut donc prétendre décrire des processus similaires quant à la spontanéité avec laquelle l'information se présente.

2.3.2- Isoler l'identité du contexte autorisant des représentations spontanées

Deuxièmement, le contexte qui permet l'immunité à l'erreur d'identification peut être précisé de manière à justifier spontanément tout un ensemble de connaissances similaires à la mémoire normale, sauf celle qui concerne l'identité du moi à travers le temps. Autrement dit, la spontanéité avec laquelle nous présumons de notre identité à travers le temps peut disparaître sans perdre celle avec laquelle nous présumons du reste. Reconnaître la vulnérabilité à l'erreur de mauvaise identification ne mène donc pas à une altération massive de la pensée, mais à une modification mineure.

¹³⁴ R. J. Howell (2007), p.588-589.

¹³⁵ C. Rovane reprend cette stratégie pour répondre à J. McDowell. « In first-person thought, a person engages in a feat of self-reflexive reference. (...) There is a standard analysis that brings out very well just how this general conception comes into the meaning of "I", viz. "the thinker of this thought." The reflexive aspect of "I" is taken care by the pronoun "this" which refers to the same thought in which it figures. (...) this analysis becomes "the series of psychologically related intentional episodes to which this one belongs.", C. Rovane (1993), p.86

2.3.2.1- Immunité à l'erreur d'identification comme absence d'effort justifié

Dans la mémoire, nous nous représentons directement, sans effort et sans intermédiaire une expérience passée en étant pleinement autorisé d'y croire. Bref, il n'y a nul besoin de faire entrer un raisonnement dans le flux de conscience pour venir justifier nos assertions relativement à ce qui s'y est passé. Simplement rapporter les détails de l'expérience suffit. L'identité du moi intervient alors comme critère autorisant l'utilisation de ce mode automatique de représentation. En revanche, si je me représente la vie des autres aussi spontanément, ils vont exiger de moi des justifications supplémentaires. Je n'aurai donc pas aussi directement autorité pour décrire leur vie. Voilà le sens de la spontanéité mnésique¹³⁶.

C'est dans ce contexte que la discussion autour de l'immunité à l'erreur d'identification prend son sens. Selon G. Evan, J. McDowell, T. Burge et A. Hamilton, il est impossible de concevoir des cas *même imaginaires* où je suis autorisé de me représenter ainsi les expériences d'autres personnes. Ils s'appuient sur le sentiment de bizarrerie qu'accompagne l'énoncé « je suis sûr et certain que durant ce voyage à Venise, les choses se sont passées de telle et telle manière ; mais je me demande seulement si c'est arrivé à moi ou à quelqu'un d'autre ». Notre réflexe serait d'exiger une justification supplémentaire sur la façon dont la personne le sait. Par témoignage ? Par lecture ? Si elle n'est pas certaine que ce soit arrivé à elle, nous ne lui concéderions pas cet énoncé comme si elle rapportait un souvenir. C'est pourquoi cette représentation ne peut pas avoir la forme de la mémoire et simplement laisser ouverte avec possibilité d'erreur la question « à qui cela est arrivée ? »

Cependant C. Rovane, A. Colliva, J. Pryor et C. Buford apportent des arguments contre l'immunité à l'erreur d'identification qui sont éclairantes à cet effet.

¹³⁶ Ce qui résume les réflexions de J. McDowell (1997) et de T. Burge (2003).

L'immunité à l'erreur d'identification peut signifier simplement qu'en l'absence d'un effort conscient de m'identifier à moi-même via la mémoire, je ne peux pas me tromper. Pour se tromper, il faut tenter quelque chose et je ne tente rien. A. Collivia résume ainsi cette immunité : « A judgment of the form « a is F » is immune to error through misidentification *relative to a subject's rational grounds* if and only if the subject's own rational grounds do not contain any identification component »¹³⁷. Dans ce cas, l'être qui résulte d'une duplication de mon esprit par double transplantation d'hémisphères cérébraux ne se trompera pas en s'attribuant mes expériences, car il ne m'aura jamais visé comme cible de référence, pour ensuite commettre l'erreur de me confondre avec lui.

2.3.2.2- Vulnérabilité à l'erreur d'identification qui préserve l'autorité du contexte

Cependant, cela n'empêche pas d'être vulnérable à une erreur d'identification entendue dans un sens que J. Pryor qualifie de « quel objet ». Il y a un ensemble de raisons présentes dans le contexte qui m'autorise à me représenter spontanément les expériences passées comme étant arrivées et à croire qu'elles sont les miennes. Ce ne sont pas des raisons dont j'ai forcément conscience. Une information additionnelle vient me montrer qu'elles ont été vécues en fait par une autre personne sans défaire pour autant la connaissance que j'en ai¹³⁸. Voici une manière d'appliquer ce raisonnement à celui de D. Parfit. La vivacité, la richesse de détail et toutes les autres caractéristiques apparentes de ma représentation combinées à la présence d'un phénomène neurologique x (exemple : une « trace mnésique ») coïncident avec un haut degré de fiabilité. Elles m'autorisent habituellement à mémoriser les

¹³⁷ A. Collivia (2006), p.420. Elle cherche à préciser en fait le concept d'une immunité à l'erreur d'identification « de re » avancé par J. Pryor (1999), p.273-280.

¹³⁸ Il formalise cette erreur ainsi : « When that proposition is believed on grounds G, it is possible for those grounds to be defeated by undercutting evidence in such a way that the following two conditions hold : i) the combinaison of G and the undercutting evidence no longer justifies you in believing of a that it is F ; but ii) by itself, offer you knowledge that $\exists x Fx$. », J. Pryor (1999), p.284. A. Collivia (2006) et C. Buford (2009) usent d'une notion de vulnérabilité à l'erreur relative au contexte qui s'en inspire.

expériences passées, mais également à croire automatiquement que je suis celui qui les a vécus et que je suis ainsi la cause de cette représentation. Je ne me réfère pas à quelqu'un dans le passé pour m'identifier à lui, mais j'utilise sans m'en rendre compte ces raisons issues du contexte pour croire en ma présence dans le passé. Je peux commettre une erreur d'identification, car ces raisons m'autorisent peut-être seulement à me représenter de façon fiable et spontanée l'expérience passée, mais non à croire en ma présence dans le décor¹³⁹.

La fiabilité de la mémoire et notre présence dans l'expérience passée vont habituellement ensemble et c'est pourquoi personne ne se donne la peine de s'assurer que c'était bien eux et non quelqu'un d'autre qui étaient présents après s'être représenté vivement, précisément, richement, spontanément et avec conviction une expérience passée via un processus neurologique fiable. C'est ce qui explique la présence d'une immunité liée à une absence d'effort conscient d'identification. Cependant, cela n'empêche pas que des événements étranges puissent venir séparer cette association automatique¹⁴⁰.

Selon J. Pryor, l'erreur de mauvaise identification peut avoir la même forme que celle concernant une silhouette au loin que nous identifions erronément à une personne. Les raisons de penser qu'une personne connue se trouve là sont aussi des raisons de croire qu'une personne s'y trouve. Nous commettons une erreur d'identification si des informations additionnelles viennent invalider son identité, mais non le fait qu'il y en a quelqu'un. La raison invalidante ne falsifie pas l'ensemble de la connaissance, mais seulement la partie identificatoire¹⁴¹. La duplication de notre esprit par double transplantation d'hémisphères cérébraux dans deux corps similaires au nôtre peut produire des erreurs d'identification de même nature. C. Burford l'illustre de la façon

¹³⁹ J. Pryor applique ce raisonnement à l'argument de la fission avancée par Parfit. J. Pryor (1999), p.293.

¹⁴⁰ C'est l'explication commune donnée à ce genre d'immunité par C. Rovane (1993), p.90-93 et J. Pryor (1999), p.290-292.

¹⁴¹ J. Pryor (1999), p.296-297.

suivante. J'ai enterré un trésor et l'être qui résulte de la fission va le chercher. Intuitivement, il sera autorisé de croire qu'une personne a mis un trésor là via le même processus que la mémoire et la nouvelle de la duplication ne va qu'invalider le fait que ce soit lui qui l'a mis là¹⁴².

2.3.2.3- Effet de la généralisation

Ajoutons que si ce genre de scénario se produit régulièrement, l'ensemble du processus mnésique ne sera pas invalidé comme devant être complété d'enquêtes sur le monde extérieur, mais il ne s'accompagnera simplement pas de la croyance automatique en une identité¹⁴³. Autrement dit, la représentation du passé via cette cause continuera de justifier aussi spontanément des informations, mais seulement la présence d'un même être ayant traversé ces événements ne sera plus prise pour acquise. La perte de spontanéité sera donc restreinte à la représentation de soi dans le passé, mais non à la représentation de ce passé¹⁴⁴. Dans ce cadre, asserter avec conviction des détails sur une expérience passée, puis questionner ensuite la nature de l'identité des personnes impliquées ne nous disposerait plus à exiger des justifications supplémentaires venant de l'extérieur.

Voici une comparaison illustrant cette idée. « J'ai joué aux cartes de telle ou telle manière, mais je ne sais tout simplement pas si c'était un Mardi ou un Mercredi » ne nous dispose pas à exiger de l'énonciateur des justifications venant du monde extérieur sur le pourquoi il croit avoir joué aux cartes de cette manière. Le témoignage de sa mémoire continue d'avoir spontanément autorité malgré une indécision relativement à la localisation temporelle de l'événement. Dans le même ordre d'idée, si les scénarios de D. Parfit se généralisaient, l'énoncé « Une telle partie

¹⁴² C. Buford (2009), p.476-477.

¹⁴³ R. Roache répond à J. McDowell par un argument similaire dans R. Roache (2006), p.351-352.

¹⁴⁴ C. Rovane (1990) suggère également en ce sens que la perte de spontanéité imaginée dans la fission n'est que le résultat de la surprise. Si ce genre d'événements étaient su à l'avance, les gens se la représenteraient intérieurement avec la même spontanéité que la mémoire ou l'anticipation normale.

de carte a eu lieu de telle ou telle manière, mais j'ignore si c'est moi, un pré-dupliqué ou un de mes donateurs de traces mnésiques qui l'a joué » aurait la même autorité spontanée que la mémoire. L'indécision relativement au sujet de l'expérience sera aussi peu invalidante contextuellement que sa localisation temporelle.

De la même manière, si les gens savaient à l'avance qu'ils reçoivent des traces mnésiques d'autres personnes à la manière de Jane et Paul et que les représentations qui en découlent sont rarement contredites, ils pourraient penser des choses comme « Un tel voyage a eu lieu de telle et telle manière, mais est-ce arrivé à moi ou à Paul¹⁴⁵ ? Ah, je le quasi-mémorise mieux maintenant, c'est arrivé à Paul parce que la dernière transplantation de ses traces avait eu lieu juste après son voyage ! » Dans ce cas, « Un tel voyage a eu lieu de telle ou telle manière » restera une représentation ayant une autorité spontanée qui ne nécessite pas une confirmation du monde extérieur. Pourquoi ? Parce que les gens assumeraient aussi automatiquement la présence d'une causalité fiable de la représentation dans cette fiction qu'ils assument automatiquement leur présence passée dans la mémorisation actuellement. Cette justification serait reléguée dans le contexte externe au flux de conscience parmi le bassin de connaissances prises pour acquises qui sont censées définir la perspective du moi selon T. Burge. Donc, l'identité du moi à travers le temps est en fait superflue pour comprendre cet aspect de la mémorisation.

2.3.3- Support aux quasi-mémoires

En résumé, D. Parfit a décrit un point de vue quasiment similaire à la mémoire sans se référer à la présence d'un moi identique afin de surmonter un problème de circularité, de clarifier le sens de son réductionnisme et de montrer la présence d'une croyance illusoire en une chose de plus que le corps et un ensemble de processus

¹⁴⁵ D. Parfit décrit la situation comme si Jane enquêtait dans le monde pour savoir que sa représentation porte sur le voyage de Paul, mais, comme le fait remarquer J. Pryor (1999), p.296, elle pourrait très bien avoir reçu l'information préalablement de manière à ne s'appuyer que sur sa représentation pour tirer cette conclusion.

mentaux. Selon G. Evan, T. Burge, J. McDowell et A. Hamilton, la perspective du moi s'exprime à travers la spontanéité avec laquelle des informations à propos du passé peuvent être utilisées dans un contexte. Tenter comme D. Parfit de la définir comme une conclusion à partir de données impersonnelles préalablement comprises la déforme. Les arguments précédents suggèrent plutôt que cette spontanéité de la représentation n'a rien de spécifique à la perspective du moi et existe à la troisième personne. D'autre part, l'automatisme en vertu duquel des informations fiables à propos du passé sont utilisées peut exprimer autre chose que la référence à un moi. Il peut désigner la manière dont les événements mentaux s'enchaînent à l'intérieur de « ceci », se transmettent entre des sujets vaguement définis ou se passent d'une personne à l'autre. Des scénarios étranges peuvent me prouver que je me représente le passé d'un autre sans invalider l'autorité avec laquelle je le rapporte. Il suffit d'un contexte qui en normalise l'occurrence. Nous pouvons décrire alors un phénomène quasiment similaire à la mémoire normale sans référence à la présence d'un moi identique à travers le temps.

2.4- Présence de l'illusion cartésienne

Cette discussion clarifie un aspect de la réflexion de D. Parfit discuté au chapitre précédent. Selon lui, la perspective du moi dans la mémoire a une composante illusoire : elle s'accompagne de la croyance implicite que nous sommes une chose simple et inanalysable au-delà du corps, des processus mentaux et du langage pour en parler. La mémoire est ainsi une faculté partiellement trompeuse. Le réductionnisme ne rend donc pas compte de nos intuitions, mais les défie partiellement. Les objections de G. Evan, J. McDowell, T. Burge et A. Hamilton visent plutôt à décrire la perspective à la première personne sans faire intervenir une telle théorie de l'erreur. Selon eux, la présence d'un contexte permettant de se représenter des expériences passées directement sans formuler d'hypothèses identitaires rend compte de l'autorité associée à notre perspective courante. C'est en sortant de ce contexte que des

croyances non réductionnistes de type cartésiennes apparaissent. Or, cette sortie ne nous informe pas sur ce en quoi consistent le moi et son point de vue.

Dans ce qui précède, c'est la possibilité de concevoir la spontanéité de la mémoire sans se référer à la présence d'un moi identique à travers le temps qui était en débat. Celui-ci apparaît parce que la présence d'un même être sert de critère habituellement pour établir la correspondance entre l'expérience représentée et la réalité sans passer par une observation du monde extérieur. Je m'imagine un passé. Parce qu'il est arrivé à moi et non à une autre personne, j'assume automatiquement qu'il a eu lieu de cette manière. Cette supposition est reconnue par mon entourage comme ayant autorité.

2.4.1- Autorité à la première personne et non réductionnisme

L'intuition non réductionniste se nourrit de cette dimension automatique de la pensée, car toute analyse de ce en quoi consiste le moi semble la détruire. C'est le sens du malin-génie de Descartes. Si je suis un corps, je peux douter de ma propre existence (présente ou passée), car son existence est médiatisée par des perceptions pouvant être trompeuses. Or, je constate de façon certaine ma présence par une voie interne : la pensée. Je dois donc être une chose purement pensante. Autrement dit, le *pouvoir* de s'assurer de sa propre existence sans passer par l'intermédiaire d'observations faillibles à propos du monde extérieur montre que je suis plus qu'une chose étendue dans l'espace. En effet, une telle chose n'a pas ce *pouvoir*¹⁴⁶. L'objection adressée par J. Butler à l'encontre de Locke selon quoi la mémoire ne peut pas être l'identité du moi, car il faut d'abord établir l'identité du moi pour déterminer que je mémorise peut être interprétée en ce sens¹⁴⁷. Cette apparente conscience du passé demeure potentiellement trompeuse tant que nous n'avons pas d'abord établi qu'elle porte sur

¹⁴⁶ Descartes (1979), méditation I, II et VI.

¹⁴⁷ J. Butler (1736), p.302. Le sens qu'il donne à cette objection demeure vague, donc ce qui suit en constitue seulement interprétation possible.

mon passé et non celui d'un autre. L'impression de mémoriser constitue, comme le corps chez Descartes, une évidence faillible en faveur de ma présence dans le passé. La mémoire ne peut donc rendre compte de l'infailibilité de ma présence. C'est plutôt l'introspection infailible de ma présence passée qui donne à la mémoire la forme d'une certitude.

Cette réflexion ne s'appuie pas seulement sur l'imaginaire tordu du malin-génie, mais aussi sur le contexte de la vie courante. La mémoire est largement conçue comme fournissant un accès infailible à des faits à propos du passé. Ce pouvoir semble marquer de façon claire et tranchée la différence entre les perspectives à la première et à la troisième personne. L'idée que nous accédons exclusivement à l'effet que cela faisait pour nous de vivre telle et telle expérience et donc que nos esprits sont radicalement séparés les uns des autres par une dimension privée est reconnue comme allant de soi¹⁴⁸. De plus, le langage pour parler de la mémoire suggère qu'elle constitue une *vision de soi* dans le passé. Nous disons que nous avons une image claire, fidèle, vive et détaillée de nous qui est organisée de façon géométrique. Nous *regardons* notre passé sous un *angle* et le tout est articulé autour d'un *point de vue* que nous *occupons*. Nous n'avons pas l'impression de reconstruire la scène après coup en enquêteur, mais d'y accéder directement. Ces observations font partie du contexte d'utilisation normale de la mémoire et d'un bassin de croyances largement répandues dans la psychologie populaire. Le non réductionnisme intervient comme une manière d'expliquer ces apparences en s'appuyant sur l'intuition qu'aucune analyse du moi en termes de corps, de cerveau ou de processus mentaux particuliers ne peut en rendre compte sans en faire au final une sorte d'illusion¹⁴⁹.

¹⁴⁸ Voir T. Nagel (1981) pour une analyse de cette intuition.

¹⁴⁹ Je m'inspire ici largement de la description que D. C. Dennett (1991) fait de la psychologie du sens commun entourant la conscience et de son explication du cartésianisme.

2.4.2- Dimension contre-intuitive du réductionnisme et de ses critiques

Effectivement, ni l'approche de D. Parfit, ni les critiques présentées dans ce chapitre ne satisfont cette intuition. Pourquoi ? Parce qu'elles ont en commun d'expliquer la spontanéité avec laquelle la mémoire représente le passé comme *résultant de données externes au flux de conscience*.

Chez D. Parfit, elle est établie via la causalité physique. Si Jane peut se représenter de façon fiable le passé de Paul ou si les êtres qui suivent ma duplication peuvent le faire de mes expériences, c'est parce que leurs représentations sont causées par les mêmes circuits neurologiques. Ce n'est donc qu'en s'appuyant sur un mécanisme externe que de tels êtres peuvent se représenter le passé comme une façon de connaître aussi spontanée que la mémoire.

G. Evan, J. McDowell, T. Burge et A. Hamilton objectent à D. Parfit cette conséquence, mais ne proposent rien de mieux. Ils lui reprochent d'exiger des êtres mis en scène dans ses fictions un effort réflexif de questionnement qui déforme l'apparence de la mémoire, mais font reposer autant que lui cette spontanéité de la mémoire sur une donnée extérieure au flux de conscience : le contexte.

Ils ne nient ainsi pas la faillibilité de la mémoire et la possibilité que je puisse me tromper sur mon existence passée, mais affirment simplement que nous devons y être nonchalants pour donner à nos états mentaux l'*apparence* de la mémoire. C'est le sens de l'immunité à l'erreur d'identification qu'ils invoquent à l'encontre de D. Parfit. Or, ce n'est pas l'exigence de se poser quelques questions sur la fiabilité des traces mnésiques, les conséquences d'une double transplantation de cerveau ou l'usage du mot « moi » ou « un autre » qui rendent les descriptions impersonnelles de la mémoire si différentes de la mémoire courante, mais l'intuition que la mémoire devient alors un processus pouvant être faillible, car il n'existe aucune voie d'accès *interne* qui nous relie au passé. Leurs critiques aboutissent tout autant à cette

conclusion contre-intuitive que D. Parfit. S'il ne s'agit que d'être autorisé par le contexte à être nonchalant quant à certaines possibilités compromettantes afin de tirer une conclusion, il n'est pas difficile d'imaginer comment ces conditions peuvent relier plusieurs êtres. Ce qui est plus difficile, c'est de continuer à croire que l'identité du moi à travers le temps est un mode d'accès direct au passé.

La vulnérabilité à l'erreur d'identification décrite précédemment défie cette intuition. En effet, la mémoire n'est pas comprise comme une vision directe de soi dans le passé, mais comme l'utilisation d'une source d'informations assez fiable pour ne plus être remis en cause. Le cerveau est alors l'équivalent d'un disque dur qui enregistre de l'information. La mémoire ne nous relie pas directement au passé, mais passe par son intermédiaire. Ce qui explique l'apparence d'un lien direct, c'est que nous n'avons pas à constamment questionner la fiabilité de ce médium. C'est parce que la mémoire est ainsi comprise que nous pouvons concevoir des cas où ce mécanisme continue de fonctionner entre plusieurs êtres. Les personnes impliquées peuvent alors se tromper sur l'identité de ceux qui ont vécu les expériences représentées sans douter du fait qu'elles ont été vécues.

2.4.3- Dissolution de la séparation radicale entre deux perspectives

À la lumière de ces précisions, les arguments de Parfit décrivant un processus similaire à la mémoire sans la référence à un moi à travers le temps peuvent être analysés comme une stratégie défiant l'intuition non réductionniste d'une séparation radicale entre la manière dont je me représente fidèlement mon passé et celui des autres. Se le représenter « de l'intérieur » ou « de l'extérieur » devient aussi trivial que de se le représenter « à l'envers » ou « à l'endroit ».

Les différentes fictions qu'il présente joue le rôle de différentes teintes de gris dans un spectre conçu préalablement en termes de noir et blanc. Au milieu, la distinction se réduit à un choix de perspective. Habituellement, dès qu'une expérience est

arrivée à moi plutôt qu'à une autre personne, elle m'apparaît avec un degré de fiabilité incomparable. Elle est directe plutôt qu'indirecte, intérieure plutôt qu'extérieure, immédiatement plutôt que médiatement confirmée. C'est noir ou blanc. Cependant, face à l'éventualité d'une duplication de mon esprit par double transplantation d'hémisphères, j'imagine plus facilement une zone de gris où ces distinctions sont deux façons de voir la même chose, comme l'image classique du canard-lapin. Le degré de fiabilité, le direct ou l'indirect, l'intériorité ou l'extériorité de l'état mental qui reliera l'être futur et mes expériences présentes ne diffèrent pas selon que je conçois cet être comme étant moi ou une autre personne. Je suis en présence d'une zone de gris où ces distinctions n'influencent pas ma capacité de comprendre la spontanéité de l'état mental. L'intuition cartésienne d'une séparation radicale entre le point de vue à la première et la troisième personne disparaît.

Lorsque D. Parfit inclut les pensées dans « ceci » plutôt que de les attribuer à « moi », il vise le même objectif. « Ceci » apparaît clairement comme une entité flexible dont les frontières peuvent varier selon un choix de perspective.

Si cette conclusion a une application dans la réalité courante, c'est de montrer comment la frontière qui sépare nos esprits n'est pas dans la *manière* dont nous nous représentons nos vies par rapport à celles des autres, mais dans la *quantité* de représentations que nous en avons¹⁵⁰. Les objections mentionnées dans ce chapitre cherchaient à préserver la radicalité de cette distinction entre la perspective du moi sur lui-même et la perspective du moi sur une autre personne en invoquant la spontanéité de la représentation comme marqueur. Elles s'opposent au non réductionnisme en attribuant au contexte plutôt qu'à une chose spéciale le pouvoir de

¹⁵⁰ Ce qui expliquerait cette citation : « When I believed that my existence was such a further fact, I seemed imprisoned in myself. My life seemed like a glass tunnel, through which I was moving faster every year, and, at the end, there was darkness. When I changed my view, the walls of my glass tunnel disappeared. I now live in the open air. There is still a difference between my life and the lives of other people. But the difference is less. » D. Parfit (1984), p.281.

marquer cette distinction. Les arguments précédents tendent à montrer qu'elles échouent à le faire autant qu'une approche cartésienne.

2.5- Résumé

Nous avons vu comment réduire l'existence du moi en l'existence d'un processus mental comme la mémoire posait un problème de circularité. Pour reconnaître un état mental comme étant une mémoire, il semble que nous devons d'abord avoir établi qu'il relie un même être à ses propres expériences et donc que le terme à analyser, l'identité du moi, se retrouve dans les éléments de l'analyse, la mémoire. Afin de surmonter ce problème, D. Parfit a décrit un point de vue quasiment similaire à la mémoire qui ne fait pas référence à l'identité d'un moi. Il s'agit d'une représentation dont l'apparence est similaire, mais dont la relation à l'expérience passée est établie plutôt par la causalité. Il présente des scénarios de transplantation de traces mnésiques entre deux personnes, de fission ou d'un univers alternatif similaire où les expériences sont incluses dans des séquences désignées par « ceci » pour rendre ce point de vue alternatif. Par ce détour, il vise également à clarifier le sens du réductionnisme en tant que thèse selon quoi notre existence, par rapport à ce point de vue, n'ajoute qu'un langage de plus. Or, notre intuition serait qu'une chose de plus manque et celle-ci est illusoire.

Plusieurs auteurs lui ont reproché de déformer radicalement le point de vue à l'œuvre dans la mémoire. Il en fait un état mental où l'identité est vulnérable à une erreur d'identification. Il lui donne la forme d'une illusion ou d'une hypothèse devant être confirmée par le monde extérieur. Il ne reconnaît pas comment la perspective du moi exprime la présence d'un ensemble de connaissances préalables permettant l'action. Bref, son point de vue altère la spontanéité à l'œuvre dans la mémoire. Or, le moi se caractérise par le contexte qui la permet. En sortir par le détour de l'imagination ne nous informe donc en rien de sa nature véritable.

Il est possible de trouver dans les objections soulevées contre l'immunité à l'erreur d'identification des moyens de répondre à ces objections. Ce genre d'immunité n'est pas spécifique à la première personne et existe dans des pronoms comme « ceci ». Ensuite, il est possible de préserver un contexte autorisant la spontanéité de la représentation mnésique en rendant vulnérable simplement l'identité du moi à travers le temps. Seulement l'invraisemblance des raisons nous permettant d'en douter nous dispense de le faire. Finalement, ni les approches centrées sur le contexte normal, ni celle de Parfit substituant l'identité à la causalité comme moyen d'établir un lien entre la mémoire et l'expérience passée ne parviennent à répondre à l'intuition d'une distinction radicale entre la perspective à la première et à la troisième personne, car elles font toutes les deux reposer la spontanéité de la représentation dans une entité externe au flux de conscience. L'argumentation de D. Parfit agit comme moyen de défier cette intuition, tout en reconnaissant sa présence.

CHAPITRE III

LE CONTEXTE DE NARRATION

Plusieurs auteurs reprochent à D. Parfit de négliger l'intégration des expériences conscientes dans un récit de vie. Ils comparent le moi à un texte dont les mots expriment l'ensemble. Suivant cette logique, l'expérience consciente ne peut être décrite sans référence à un moi identique à travers le temps, car elle reçoit sa signification du récit à l'intérieur duquel elle s'inscrit et celui-ci est l'histoire d'un être unique et individuel. En cherchant à la reformuler autrement, D. Parfit commet l'erreur d'en faire quelque chose de différent ou d'incompréhensible.

J'introduirai ce problème en le plaçant dans son contexte. Comme mentionné dans le chapitre précédent, D. Parfit cherche à éviter un problème de circularité. Il définit un concept de quasi-mémoire et de quasi-intention où les représentations entretiennent une relation causale et phénoménale similaire à la représentation normale du passé et des actions futures, mais qui peuvent logiquement se transmettre d'une personne à une autre. Il met en scène des fictions rendant concevables de tels processus. Il en ressort apparemment une conception morcelée, passive et brève des expériences conscientes.

Je présenterai par la suite des objections selon lesquelles l'intégration dans un tout, la dimension active et la durabilité des expériences conscientes sont des conditions pour leur donner un sens. Je les opposerai ensuite à trois réponses : l'existence de systèmes séparés dans le cerveau contredit cette intégration incontournable de l'expérience dans la vie d'un moi, ces conditions sont contingentes et insuffisantes, puis nous pouvons les décrire en substituant au « je » des pronoms démonstratifs autoréférentiels différents.

Je synthétiserai cette discussion par deux conclusions. Premièrement, certaines observations courantes facilement accessibles suffisent à défendre D. Parfit sans

spéculer sur le fonctionnement du cerveau ou l'usage de fictions. Deuxièmement, décrire de façon légèrement différente une seule et même réalité, puis multiplier les contextes de signification ne posent aucun problème en général. Il est donc étrange d'en faire une exception appliquée à l'identité du moi. La présence d'une croyance non réductionniste l'explique de façon crédible.

3.1- Approche de D. Parfit

D. Parfit analyse l'existence du moi à travers le temps au moyen de processus comme la mémoire, la poursuite d'intention et la persistance de traits de caractère et de croyances entretenant une relation causale avec le corps. Je ne suis pas une chose qui transcende ces composantes et le langage pour en parler. Or, cette analyse pose un problème de circularité. Pour reconnaître un état mental comme étant une mémoire plutôt qu'autre chose, il faut qu'il me relie à moi vivant une expérience passée. De la même manière, pour le reconnaître comme étant une intention, il doit porter sur moi performant une action présente ou future. Les termes de l'analyse de D. Parfit contiennent donc ce qu'il y a à analyser : le moi. Il a surmonté cette difficulté en concevant des processus quasiment à la première personne qui ne font pas référence à l'identité d'un moi à travers le temps pour être compris¹⁵¹.

Dans le chapitre précédent, nous avons discuté de ce problème sous l'angle de l'autorité de la connaissance à la première personne. Nous en discuterons ici, mais à propos de la *signification* des expériences conscientes. De nombreux auteurs lui reprochent de les réduire à des phénomènes atomisés, passifs et éphémères dont le contenu est compris hors contexte. En réalité, les expériences conscientes signifient quelque chose en vertu du rôle durable qu'elles jouent dans une vie voulue par un agent. D. Parfit ne peut éviter la circularité qu'en modifiant radicalement cette signification¹⁵².

¹⁵¹ Il résume cet effort ainsi : « There are two unities to be explained : the unity of consciousness at any time, and the unity of a whole life. These two unities cannot be explained by claiming that different experiences are had by the same person. These unities must be explained by describing the relations between these many experiences, and their relation to this person's brain. And we can refer to these experiences, and fully describe the relation between them, without claiming that these experiences are had by a person. » D. Parfit (1984), p.217.

¹⁵² Parmi ceux qui lui formulent de telles critiques, il y a S. Blackburn (1987), C. Korsgaard (1984), J. Proust (2003), A. Rudd (2005), M. Schechtman (1990, 2005), M. Slors (2001) et R. Wollheim (1984).

3.1.1- Morcellement

L'imaginaire argumentatif de Parfit morcelle l'esprit humain en le traitant comme une somme et un enchaînement de parties plus élémentaires¹⁵³. Nous devons concevoir plusieurs manières de lui en soustraire ou de lui en additionner. Puis, vient alors la question : « le résultat est-il moi ou une autre personne ? »

Il décrit ainsi comment un morceau du cerveau de Paul peut être copié et transplanté dans le cerveau de Jane pour lui donner un état mental qui ressemble à la mémoire. Il est ainsi possible de concevoir un processus similaire à la mémoire qui ne nécessite pas la référence à un seul et même être qui passe par toutes ces étapes. La mémoire normale d'un individu ne serait donc qu'une somme assez grande de tels morceaux impersonnels. Le problème de circularité est donc surmonté¹⁵⁴.

Ensuite, il imagine un spectre de décomposition et de recomposition du soi. Un scientifique a le pouvoir de combiner un certain pourcentage de nos états mentaux avec un certain pourcentage de ceux de Greta Garbo. Parfit suppose que toutes les variations de pourcentage ne se distinguent entre elles que de changements mineurs¹⁵⁵. Ainsi, entre l'être résultant qui a 45 % de nous et 55 % de Greta Garbo et celui qui a 46 % de nous et 54 % de Greta Garbo, la seule différence repose sur quelques états mentaux de plus de nous et quelques états mentaux de moins d'elle. C'est de ce point de vue qu'il en conclut qu'au milieu du spectre, il y aurait une zone floue où l'être résultant pourrait être nous ou Greta Garbo en fonction d'un choix de mots et non en fonction d'un fait. Il use de cette stratégie de visualisation pour faire ressortir comment le moi peut être l'addition d'un langage sur une réalité

¹⁵³ Cette dimension est manifeste dans D. Parfit (1984), p.205-206 et p.222. Il traite alors l'esprit comme une *addition* et un enchaînement de souvenirs, d'intentions, de traits de caractères et de croyances dans le temps.

¹⁵⁴ D. Parfit (1984), p.219-223.

¹⁵⁵ Il affirme que la dépendance des états mentaux envers des états du cerveau justifie de penser que si de telles opérations se produisaient, il n'y aurait pas un endroit sur le spectre où un changement mineur provoquerait une altération majeure de la psychologie d'une personne. D. Parfit (1984), p.237-238.

préalablement comprise sans référence à lui¹⁵⁶. Il s'agit donc d'un deuxième argument surmontant le problème de circularité.

Son argument central est un scénario de double transplantation d'hémisphères cérébraux équipotents dans des corps similaires au mien, ce qui est appelé *la fission*. Dans ce cas-ci, nous devons prendre pour acquis que l'ensemble des parties de notre esprit serait poursuivi à deux endroits différents et donc que deux êtres auraient des raisons égales d'être nous¹⁵⁷. Ce qui est soustrait ici comme ingrédient de notre esprit est la singularité du flux de conscience qui nous relie au futur ou au passé. Cette fiction a la même structure que les deux précédentes. Un seul élément est modifié, la singularité du flux de conscience, mais tout le reste est censé rester pareil¹⁵⁸. D. Parfit traite donc ici encore l'esprit comme quelque chose qui se découpe en partie isolable de l'ensemble. Les conclusions qu'il tire de cette situation sont doubles. Premièrement, considérer le résultat comme étant nous à deux places ou comme étant deux êtres qui poursuivent notre vie est un choix de mots. Nous pouvons donc comprendre entièrement la nature du processus mental impliqué sans décider de nous l'attribuer, contournant ainsi la circularité¹⁵⁹. Deuxièmement, si le résultat décrit est l'apparition de deux êtres, alors il montre comment d'autres personnes peuvent quasiment mémoriser nos expériences, performer nos intentions, poursuivre nos traits de caractère et nos croyances d'une manière indiscernable de nous. Seulement

¹⁵⁶ D. Parfit (1984), p.236-243 et (1995), p.21-25.

¹⁵⁷ « My body is fatally injured, as are the brains of my two brothers. My brains is divided, and each half is successfully transplanted into the body of one of my brothers. Each of the resulting people believes that he is me, seems to remember living my life, has my character, and is in every other way psychologically continuous with me. And he has a body that is very like mine. » D. Parfit (1984), p.255.

¹⁵⁸ Bien sûr, la transplantation de cerveau et la division des hémisphères sont également des différences, mais D. Parfit les présente comme ne faisant aucune différence sur l'apparence des états mentaux subséquents et donc comme étant triviales. Le changement majeur sur lequel il veut mettre les projecteurs, c'est la duplication du flux de conscience à partir d'une partie essentielle de notre corps. Voir D. Parfit (1984), p.255.

¹⁵⁹ D. Parfit (1984), p.259-260.

l'étiquette verbale « même être » sera soustrait, mais tout le reste sera similaire. Donc, de tels processus mentaux peuvent être analysés sans la notion de moi¹⁶⁰.

Finalement, D. Parfit conçoit des êtres qui sont similaires à nous avec la seule différence qu'ils incluent les événements mentaux dans des séquences désignées par « ceci » au lieu de les traiter comme des états appartenant à une chose indiquée par le pronom « je ». Ainsi, au lieu d'affirmer « Je mémorise telle expérience », l'expression « Ceci inclut la mémorisation de telle expérience » aurait lieu¹⁶¹. Le morcellement de l'esprit est encore présent de deux manières. D'une part, la notion de séquence suppose que l'esprit peut être conçu comme une succession de segments distincts les uns des autres. Ensuite, il s'agit d'effectuer un raisonnement ayant la forme « imaginons que tout reste pareil, sauf... » Il faut isoler une variable d'un ensemble sans en affecter l'apparence. Dans ce cas-ci, la variable, c'est le langage pour parler des expériences conscientes.

La nature du moi est ainsi analysée comme l'ajout d'un langage sur un ensemble de parties plus fondamentales préalablement comprises sans se référer à lui. C'est ce que signifie chez D. Parfit l'affirmation selon quoi le moi « consiste » en l'existence d'un corps et d'un ensemble de processus mentaux inter-reliés et non en une chose de plus. La forme de son argumentation suggère une atomisation de l'expérience consciente. Celle-ci est conçue comme une entité ponctuelle dont la réalité est détachable du tout à l'intérieur duquel elle s'inscrit.

¹⁶⁰ D. Parfit (1984), p.260-261 et (1999), p.242-251.

¹⁶¹ C'est la thèse défendue dans D. Parfit (1999).

3.1.2- Passivité

Outre ce morcellement de l'esprit, plusieurs auteurs ont reproché à D. Parfit la passivité avec laquelle il présente l'expérience humaine¹⁶².

Sans la présence d'un moi, le mental est une succession d'événements reliés ensemble par des causes¹⁶³. Lorsque Jane quasi-mémorise l'expérience de Paul, elle ne se la représente pas comme l'ingrédient d'un projet durable qu'elle mène à terme, mais comme une simple réception d'information sans impact sur son agir à elle¹⁶⁴. La décomposition et recomposition de soi, ainsi que la fission apparaissent comme des opérations que nous subissons¹⁶⁵. De plus, la projection d'action future du point de vue de celui qui anticipe ces opérations lui demande de ne plus concevoir cette action comme quelque chose qu'il fera, mais comme quelque chose qui aura lieu d'une certaine manière (de façon similaire à la vie courante), altérant l'impression qu'elle sera sous son contrôle¹⁶⁶. Dans ce cadre, l'intention n'est qu'une sorte de poussée qu'un micro-agent envoie dans le futur à la manière d'un bâton dans une course à relais. Finalement, la notion de séquence élimine apparemment l'idée qu'un être insuffle une direction aux événements mentaux. Le mental ne fait que s'enchaîner d'une certaine manière et rien de plus¹⁶⁷.

¹⁶² C. Korsgaard (1989), mais également S. Blackburn (1997), T. Burge (2003) et M. Schechtman (2005) lui formulent explicitement cette critique. D. Parfit ne défend pas cette thèse en tant que telle. En fait, il s'objecte à cette association dans D. Parfit (1999), p.229-230.

¹⁶³ C'est que D. Parfit (1999) défend.

¹⁶⁴ C'est ce que lui reproche T. Burge (2003), p.309-312.

¹⁶⁵ « It is, I think, significant that writers on personal identity often tell stories about mad surgeons who make changes in our memories or characters. These writers usually emphasize the fact that after the surgical intervention we are altered, we have changed. But surely part of what creates the sense of lost identity is that the person is changed by intervention, from outside. » C. Korsgaard (1989), p.123-124.

¹⁶⁶ S. Blackburn (1997), p.195-198.

¹⁶⁷ « Thoughts and experiences, we may concede, can be thought of as mere happenings. But it is hard to think that way about our decisions and our acts. This is part of our lives in which, it seems, we most clearly enter in. We inject agency » résume D. Parfit (1999), p.229 en commentaire à la critique de C. Korsgaard.

De plus, la perspective avec laquelle D. Parfit analyse le moi est celle de la compréhension et non de l'action¹⁶⁸. Il y a des événements bizarres qui se produisent, puis nous analysons ce qui se passe. Par ce détour, nous nous représentons le moi comme addition d'un langage sur une réalité préalablement *comprise*. Or, le sens du moi n'est pas quelque chose que nous analysons mais que nous faisons.

3.1.3- Brièveté de l'expérience consciente

Cette compréhension atomisée et passive de l'expérience consciente s'accompagne chez D. Parfit d'une dévaluation de la vie en général comme cadre de référence à nos raisons d'agir. Selon lui, si le « moi » consiste en l'addition d'un langage sur une somme et un enchaînement de processus mentaux comme la mémoire, la poursuite d'intention et autres, alors la diminution de leur nombre à travers le temps peut constituer une raison d'accorder plus d'importance à la tranche de vie présente qu'à celles plus éloignées dans le futur¹⁶⁹.

3.2- Le moi comme narrateur

D. Parfit évite un problème de circularité en morcelant l'expérience consciente, en la présentant de façon passive et en dévaluant la vie en général comme cadre de référence à nos raisons d'agir. Est-ce qu'il réussit ainsi à nous faire imaginer des processus conscients similaires à ceux que constituent nos vies, mais sans se référer à l'identité d'un moi à travers le temps ? Des critiques le nient en affirmant que nous utilisons les expériences conscientes ponctuelles pour donner un sens à notre vie. Vivre est un acte de longue durée et non pas une succession d'événements éphémères. Cet acte intègre plusieurs moments dans une histoire. Or, ceux qui subissent un morcellement et une recomposition du soi ne peuvent pas employer leur expérience pour raconter une narration comparable à la nôtre. Le sens de ce qui leur

¹⁶⁸ C'est l'analyse qu'en fait C. Korsgaard (1989) et S. Blackburn (1997).

¹⁶⁹ D. Parfit (1984), p.312-318.

arrive est radicalement appauvri ou altéré. Ces situations ne nous informent donc pas de processus conscients similaires à ceux de nos vies actuelles.

3.2.1- M. Schechtman

M. Schechtman perçoit chez D. Parfit une confusion entre deux questions : la ré-identification du moi comme objet et la caractérisation du moi comme sujet. Dans le premier cas, il s'agit de déterminer si deux descriptions servent à se référer à un seul et même être ou à plusieurs. Dans le second cas, nous nous demandons si tel ou tel phénomène exprime authentiquement qui nous sommes. Ce qui motive D. Parfit et les héritiers de J. Locke à ne pas considérer le moi simplement comme un corps humain agissant, mais plutôt comme une capacité de poursuivre des processus mentaux comme la mémoire et les intentions, c'est l'intuition que l'importance de notre existence réside dans cette capacité et non dans la continuité d'un objet physique. Or, selon M. Schechtman, cette importance ne réside pas dans l'acte de pointer un être à deux moments et de le reconnaître comme étant nous-mêmes en comptabilisant le nombre de processus mentaux qui les relie (le moi objet), mais dans la perception d'une expérience comme étant révélatrice de qui nous sommes (le moi sujet)¹⁷⁰.

M. Schechtman attribue aussi à D. Parfit la tentative de séparer le contenu des états mentaux de l'expérience qui lui est associée. Le « contenu » renvoie à ce sur quoi il porte. Lors du transplant de souvenirs du voyage de Paul à Venise vers Jane, les deux peuvent viser les mêmes sons et les mêmes images dans leur représentation. En ce sens, l'état mental aura un contenu similaire¹⁷¹. Cependant, celui-ci n'aura pas le même *effet* pour Jane que pour Paul, car les deux associeront la scène à des éléments

¹⁷⁰ M. Schechtman (1990), p.87-90. « The intuition we have in response to the thought experiments, I claim, come from our view of persons as subjects, but the methods of identity theorists use to turn these intuition into a criterion of identity come from a view of person as objects. » M. Schechtman (1990), p.88-89.

¹⁷¹ M. Schechtman (1990), p.79.

de leur passé respectif. Les événements de Venise pourront ainsi être vécus de façon plus familière pour l'un que pour l'autre. L'expérience de se représenter ce voyage ne sera donc pas la même. Cela sera encore plus marqué s'il s'agit d'un souvenir complexe qui implique des événements marquants de la vie d'une personne. Si celui qui le reçoit n'a pas eu d'enfant et qu'il y en a un de représenté dans sa quasi-mémoire, il lui apparaîtra avec un air d'étrangeté absent chez celui qui a vécu l'expérience. Or, les souvenirs les plus riches à constituer l'identité du moi à travers le temps sont de cette nature. Toute expérience sortie du contexte de la vie d'une personne apparaîtra donc comme un magma de perceptions difficiles à interpréter, une illusion de l'avoir vécue ou une représentation considérablement appauvrie¹⁷².

M. Schechtman croit également que ces associations s'organisent autour de grands projets unificateurs à l'intérieur duquel le moi se projette. Nous nous imaginons constamment en train d'en faire l'expérience dans le futur et orientons nos actions dans ce but. L'absence de cette planification dans une personnalité dépressive correspond à un sentiment d'amoindrissement de soi¹⁷³.

Autrement dit, la mémoire des expériences passées ne porte pas sur des informations isolées à propos du passé qui s'additionne ou se soustrait pour faire une somme, mais sur un réseau d'autres expériences qui caractérisent une vie entière. La mémoire coordonne des projets que nous imaginons ressentir activement. C'est ce qui leur donne de l'importance. Parfit a donc tort de penser que les états mentaux d'une personne puissent être découpés en morceau, transférés d'une vie à l'autre et analysés comme de simples événements qui arrivent passivement dans une séquence.

¹⁷² M. Schechtman (1990), p.80-84. R. Wollheim (1984), R. Roache (2006) et M. Slors (2001) formulent une même critique.

¹⁷³ M. Schechtman (2005). « In conceiving of ourselves as continuing as agents, we project ourselves into the future as subjects as well – imagining what will bring us satisfaction, when will be filled with regret, what is likely to tempt us away for our purpose, and so on. To experience ourselves as unified agents, we may well also have to experience ourselves as unified subjects, and vice-versa.” M. Schechtman (2005), p.21. S. Blackburn (1997) défend une thèse similaire.

3.2.2- A. Rudd

A. Rudd formule une critique similaire. Il affirme que les événements mentaux particuliers forment des patterns durables dans le temps. Ils s'articulent autour de projets à long terme. Ils suivent une trajectoire. Lorsqu'ils arrivent à l'intérieur d'une même vie, nous cherchons à les interpréter de manière à les rendre cohérents entre eux plutôt que de se satisfaire simplement de leur apparente incohérence ou diversité. Cela nous indique comment nous y cherchons la présence d'une personnalité stable. En ce sens, chacune des expériences mémorisées a la forme d'un marqueur révélant la présence d'un moi. C'est ce que signifie l'appropriation d'états mentaux¹⁷⁴.

Dans ce cadre, lorsque Jane se représente le voyage de Paul à Venise, elle n'éprouve pas ce besoin d'y voir un signe de qui elle est, d'en rendre les actions et émotions cohérentes avec les siennes. En ce sens, elle n'adopte pas une perspective similaire à la première personne. De la même manière, si l'être mélangeant des aspects de nous et de Greta Garbo ne cherche pas à relier ce qu'il vit ou ce qu'il pense à nous ou à Greta Garbo en particulier, mais se contente de comptabiliser le pourcentage de ressemblance envers l'un et l'autre, il n'adopte pas non plus une telle perspective.

3.2.3- Dimension active

Un aspect de la mémoire sur lequel met l'accent J. Proust et T. Burge, c'est qu'elle est orientée vers des actions futures et ne se réduit pas à une simple récupération passive d'information. La mémoire sert à actualiser, à coordonner et à évaluer où nous en sommes relativement à des objectifs qui s'étendent dans le temps. De ce

¹⁷⁴ A. Rudd (2005), p.422-426. « We do not think of the Self as a « bare locus » which might indifferently sustain all kind of mental states, but nor do we think of people as simply Humean or Parfitian agglomerations of more or less closely connected psychological atoms. When we get to know people, we get to know their dispositions, their characteristic way of being. And what we expect to find is a certain coherence and consistency among those traits. They form a pattern, in and through which we recognise the person whose traits they are. » A. Rudd (2005), p.424.

point de vue, la présence du moi s'exprime dans la capacité de le faire¹⁷⁵. J. Proust appuie cette affirmation sur les troubles cliniques impliquant une confusion de l'identité. Dans les troubles de la personnalité multiple, c'est l'incapacité d'agir mentalement sur des parties de soi qui pousse les patients à s'en distancier. Dans le syndrome de Cotard, c'est l'absence de motivation à agir qui donne l'impression à la personne d'être morte. Dans des cas de schizophrénie, la perception d'être envahie par des pensées étrangères est liée à l'incapacité de les contrôler¹⁷⁶.

C. Korsgaard adresse une critique similaire à Parfit. Pour mener à terme des actions durables, il faut résoudre les conflits qui surgissent entre les différentes parties de nous. Or, nous n'observons pas passivement lesquelles des parties auront le dessus. Nous décidons en fonction de principes exprimant l'autorité du tout¹⁷⁷. C'est pourquoi nous ne sommes pas troublés par des changements de personnalité amenés volontairement¹⁷⁸. Nous ne sommes pas seulement des sujets, mais sommes aussi des agents. De plus, pour avoir de la valeur, ces actions doivent s'étendre durablement dans le temps¹⁷⁹. Le moi exprime donc sa présence dans les choix de vie.

Ici encore, la critique adressée à Parfit traite les états mentaux particuliers non pas comme des éléments isolés, mais comme étant des outils visant la réalisation d'une unité d'ensemble présenté sous un mode actif. Les fictions présentées par Parfit altèrent considérablement cette perspective.

¹⁷⁵ J. Proust (2003), p.499-504 et T. Burge (2003).

¹⁷⁶ J. Proust (2003), p.504-508.

¹⁷⁷ « When you deliberate, it is as if there were something over and above all your desires, something that it is you, and that chooses which one to act on. The idea that you choose among your conflicting desires, rather than just waiting to see which one wins, suggests that you have reasons for an against them. (...) This means that there is some principle or way of choosing that your regard as expressive of yourself, and that provides reasons to regulate your choices among your desires." C. Korsgaard, p.111.

¹⁷⁸ C. Korsgaard (1989), p.123.

¹⁷⁹ C. Korsgaard, p.113.

3.2.4- Application à la fission et aux séquences désignées par « ceci »

Il ne va pas de soi que ces remarques s'appliquent à l'ensemble du raisonnement de D. Parfit, car la fission et l'inclusion des événements mentaux dans des séquences désignées par « ceci » s'accordent avec la supposition que les expériences conscientes ont une signification en vertu de leur intégration dans un contexte plus large. En effet, les deux êtres suivants la fission auraient accès à l'ensemble de mon histoire. Cela leur donnerait la capacité de poursuivre mes projets¹⁸⁰. De plus, l'intégration de la pensée exprimée dans cette phrase a lieu dans « ceci » et « ceci » désigne une vie.

M. Schechtman apporte un argument selon lequel « ceci » doit être conçu comme un agent et non comme une séquence : l'imagerie accompagnant l'action est celle d'un être qui se projette dans le futur comme étant celui qui en fera l'expérience. Le futur n'est pas perçu dans l'action comme une poussée venant de l'instant présent.

S. Blackburn formule ainsi l'objection selon laquelle, face à une duplication de notre flux de conscience, nous perdrons la capacité de s'imaginer l'effet que l'accomplissement de projets futurs aura sur nous. Nous sommes contraints de stipuler qui fera quoi et donc d'agir comme si nous donnions à des êtres extérieurs à nous des commandes. Or, cela altère significativement la perspective de l'action¹⁸¹. « It [la fission] fails to engage the unity reaction, precisely because the question of what it will be like for me is ignored »¹⁸².

La réflexion de C. Korsgaard va dans le même sens. Si « ceci » est conçu comme ayant autorité sur la pensée exprimée dans cette phrase, alors il est l'agent de cette pensée. De plus, la fission implique une perte d'autorité. Les deux produits n'ont plus besoin de résoudre leur conflit, alors le tout désigné par « moi » ne leur est plus nécessaire. Cela fait passer la perspective d'un mode actif à un mode passif.

¹⁸⁰ R. Roache (2006), p.337-347 en fait la remarque.

¹⁸¹ S. Blackburn (1997), p.190-199.

¹⁸² S. Blackburn (1997), p.195

3.2.5- Résumé

En résumé, D. Parfit affirme que le moi consiste en l'existence d'un corps et d'un ensemble de processus mentaux inter-reliés comme la mémoire, la poursuite d'intention, la persistance de traits de caractère et de croyances qui s'enchaînent graduellement au fil du temps. Cette thèse lui pose un problème de circularité, car ces processus sont définis comme des états qui relient un être à lui-même à travers le temps. Il a surmonté cette difficulté en concevant des processus quasiment à la première personne qui leur sont similaires sans se référer à l'identité d'un moi à travers le temps. Il met en scène cette perspective alternative dans des scénarios de transplants de traces mnésiques, de combinaisons partielles d'êtres, de fissions et de séquences d'événements mentaux désignés par « ceci » au lieu de « je ». Cet effort suppose à la fois que les expériences conscientes peuvent être isolées du tout à l'intérieur duquel elles sont vécues et qu'elles se réduisent à de simples événements qui s'enchaînent passivement dans une séquence.

Les objections présentées dans ce chapitre accentuent le fait que la perspective à la première personne oriente l'expérience présente vers l'unification d'un tout conçu activement : moi. Le moi s'approprie en tant que sujet et non objet des expériences. Le contenu de la mémoire ne peut être séparé de l'expérience globale qui l'entoure. Ces expériences sont unifiées par des projets durables. Elles expriment des patterns stables, une quête de cohérence et une nécessité de faire des choix à partir d'une autorité. Or, les fictions de D. Parfit déforment radicalement cette perspective. Ces objections ne supposent pas la présence d'une chose pensante de plus que le corps, les expériences et les pensées particulières, mais simplement un contexte de narration. Ce contexte doit contenir les éléments rendant possible l'expression d'un tel récit. Or, D. Parfit analyse le moi à partir de contextes trop différents. Il finit par imaginer une perspective morcelée, passive et brève de l'expérience consciente alors que celle-ci est fortement intégrée à un tout, active et durable.

3.3- Réponses aux objections centrées sur le contexte narratif

Il est possible de trouver plusieurs réponses à ces objections. Premièrement, certaines données empiriques montrent que les diverses composantes des processus mentaux sont traitées séparément. En conséquence, une perspective quasiment à la première personne sans l'être est actuelle dans quelques cas cliniques. Deuxièmement, l'intégration des expériences dans un tout qui s'étend durant une vie n'est ni nécessaire, ni suffisante pour adopter une perspective à la première personne. Finalement, certaines narrations peuvent tenir compte de l'intégration des expériences dans un tout plus large et actif en référant à des pronoms autoréférentiels différents.

3.3.1- Données empiriques

G. Northoff se porte à la défense du concept de quasi-mémoire avancé par D. Parfit en s'appuyant sur la neurologie. Selon M. Schechtman, le défaut de D. Parfit est de séparer le contenu de la mémoire de l'expérience qui lui est associée. Le cerveau fonctionne pourtant au moyen de divers systèmes de traitement d'information. Il encode l'expérience, la conserve, puis la récupère. Or, l'expérience qui entoure l'encodage diffère de celle qui entoure le retrait. Pourtant, un même contenu les unit¹⁸³. En ce sens, il n'est pas exclu que le changement de contexte à l'intérieur d'un cerveau entre deux systèmes différents ressemble au changement de contexte entre Jane et Paul. D'autre part, le cerveau a également un système de classement en mémoire épisodique, autobiographique et hétéro-biographique. La première se caractérise par une forte intégration entre l'encodage et le retrait, une forte imagerie, une puissance émotionnelle particulière, de nombreux détails spatio-temporels,

¹⁸³ « As such, either personal facts or personal events can be « encoded » and « stored ». During retrieval, these personal facts/and or events are reconstructed in “several retrieval cycles” which lead to modification of the emotional background of the past/contents by the present situation, whereas the facts/contents themselves remain unchanged. Due to such alterations in the emotional background of the unchanged “contents” the present “experience” of the past contents may differ from the past/original “experience” of the same “content”. G. Northoff (2000), p.195-196.

cognitifs et sensoriels, puis un sentiment de revivre l'expérience. Par contraste, la mémoire hétéro-biographique s'accompagne simplement d'un sentiment de familiarité¹⁸⁴. Dans ce cas, quasi-mémoriser l'expérience d'un autre implique que la présence de toutes ces caractéristiques de la mémoire autobiographique s'accompagnent d'une connaissance supplémentaire sur la source de la représentation qui empêche l'identification¹⁸⁵.

G. Northoff en illustre la possibilité à partir du processus ayant mené une femme atteinte de schizophrénie à croire être une proche d'un tsar de Russie. Celle-ci avait d'abord des informations à propos de cette proche d'un tsar sous un mode hétéro-biographique, puis a commencé graduellement à se les représenter sous un mode autobiographique avant de confondre son identité. Or, dans la phase intermédiaire, elle était en possession du genre d'état mental décrit par Parfit, c'est-à-dire une sorte de représentation non-illusoire d'une expérience passée ayant les traits de la mémoire courante sans attribution à soi¹⁸⁶.

S. Klein et S. Nichols offrent une réponse similaire aux critiques de M. Schechtman. Ils invoquent eux aussi la diversité des systèmes neurologiques. Ils présentent le cas d'une personne qui témoignait après un accident cérébral avoir une représentation de son passé possédant toutes les caractéristiques de la mémoire épisodique : vivacité, clarté, spontanéité d'apparition et précision spatio-temporelle. Cependant, elle affirmait ne pas ressentir que les événements lui sont arrivés¹⁸⁷. S. Klein et S. Nichols extrapolent alors sur la possibilité qu'un mécanisme d'attribution existe

¹⁸⁴ G. Northoff (2000), p.196-197.

¹⁸⁵ G. Northoff (2000), p.198-199.

¹⁸⁶ G. Northoff (2000), p.204-207.

¹⁸⁷ S. Klein et S. Nichols (2012), p.8-13. « R. B was able to remember particular incidents from his life accompanied by temporal, spatial, and self-referential knowledge, but he did not feel the memories he experienced belonged to him. » S. Klein et S. Nichols (2012), p.8.

séparément des autres mécanismes mémoriels et qu'il peut faillir pendant que les autres fonctionnent¹⁸⁸. C'est alors un phénomène similaire à celui décrit par Parfit.

Dans le même ordre d'idée, R. Martin mentionne des personnes qui, recouvrant d'une amnésie, se ressouvenaient morceau par morceau sans accès au contexte¹⁸⁹.

Il ne s'agit que de quelques cas, mais ils illustrent comment D. Parfit est justifié de traiter les processus conscients comme étant morcelables. Le cerveau fonctionnant au moyen de systèmes distincts, un trouble pathologique peut séparer des aspects d'une expérience consciente qui vont habituellement ensemble, défiant ainsi l'intuition d'une association nécessaire. Dans ce cas-ci, ce sont les marqueurs apparents de la mémoire (vivacité, sensation de revivre, forte imagerie, etc.) qui peuvent être présents en l'absence d'identification. De plus, le changement de contexte narratif est également présent entre l'encodage et la récupération.

3.3.2- Conditions non-nécessaires insuffisantes

Un second problème avec les objections centrées autour du contexte narratif, c'est que l'intégration des expériences dans une vie n'est ni nécessaire, ni suffisante pour être le même à travers le temps.

Elle n'est pas nécessaire, car des personnes ne se projettent pas aussi loin dans le futur et se dissocient de leur passé. L'intégration des expériences dans un tout cohérent se réfère alors à une entité plus éphémère qu'un moi personnel. D.W. Shoemaker défend ainsi que l'unité significative est un moi plus bref, qui correspond à l'intensité des liens de mémoire, d'intentions, de traits de caractère et de croyance¹⁹⁰. G. Strawson, quant à lui, défend la possibilité d'un moi épisodique qui

¹⁸⁸ « R.B. recollections during is « unowned » period can be explained in the context of the view that there is specialized neural machinery that inserts the conceptual element *self* into the agent slot of an episodic memory attribution. » S. Klein et S. Nichols (2012), p.13.

¹⁸⁹ R. Martin (1987), p.85.

¹⁹⁰ D.W Shoemaker (1999).

réfléchit peu à la signification d'ensemble de ses décisions, mais apprend et s'oriente largement en restant concentré sur l'instant présent. De plus, il peut regarder le passé et le futur sans avoir la sensation qu'ils lui appartiennent¹⁹¹.

L'intégration des expériences dans une vie n'est pas suffisante non plus, car elle ne répond pas à certains problèmes discutés par les auteurs qui s'intéressent à l'identité du moi à travers le temps. S. Bernecker objectait ainsi à M. Schechtman qu'une telle approche ne résout pas le problème posé par la réplication. Plusieurs se demandent si une copie de nous serait nous ou une autre personne. Or, les objections présentées ici n'apportent aucune contribution à ce débat. La copie poursuivrait nos projets à long-terme. Pourtant, de nombreux auteurs qui se sont posé la question nient qu'ils auraient le même genre d'attitudes envers le résultat¹⁹².

3.3.3- Pronoms démonstratifs alternatifs

En fait, ces dernières remarques montrent comment le pronom « moi » est flexible. Il peut renvoyer à une narration plus ou moins brève, se restreindre au corps ou s'étendre au-delà dans certaines fictions. Pour déterminer lequel de ces usages nous *devons* privilégier, la simple « nécessité de l'intégrer dans une narration qui lui donne un sens » ne suffit pas. Cette flexibilité des pronoms démonstratifs pose un troisième problème aux approches centrées sur le contexte de narration. La dimension holistique, active et durable des expériences conscientes peut être décrite en référant à des pronoms démonstratifs suffisamment ressemblants au « je » sans être identiques.

C. Rovane s'intéresse à rendre cohérente la dimension active du moi avec la fission en introduisant un pronom « je* » qui a la propriété, contrairement au « je » normal, d'être divisible. Il renvoie à quiconque poursuit nos processus mentaux de façon

¹⁹¹ G. Strawson (1999), p.109-110.

¹⁹² S. Bernecker (2009), p.64.

similaire à notre vie courante¹⁹³. Si la fission était un événement imprévu que nous subissions, notre perspective serait différente d'en temps normal parce que nous n'aurions rien prévu pour rendre compatible les deux nouvelles vies qui en résulteraient. La sensation de perdre une autorité sur le futur serait justifiée. Cependant, si la fission était un événement prévu à l'avance, cet obstacle serait soulevé. Nous agirions en fonction de deux ensembles de projets distincts référés à l'aide du pronom « je* »¹⁹⁴. Celui-ci fonctionnerait au moyen d'une description additionnelle. Ainsi, si nous prévoyions après la fission mener une carrière philosophique d'un côté et faire le tour du monde de l'autre, nous affirmerions « je* philosophe lira x pendant que je* voyageur fera y »¹⁹⁵. Cet usage rend compte à la fois de la dimension holistique, active et durable de la perspective à la première personne sans renvoyer au moi à proprement parler, mais à quelque chose qui est quasiment le moi.

Une autre solution est d'employer l'équivalent d'un pronom à la première personne du pluriel. M. Siderits répond ainsi à C. Korsgaard que l'agent peut être conçu comme une sorte de titre que des coalitions de processus mentaux occupent de façon successive l'une après l'autre, au même titre que « la monarchie britannique » ou « le gouvernement du Québec ». Autrement dit, le référent du pronom démonstratif servant à un acte d'autoréférence n'a pas à être conçu comme une réalité singulière¹⁹⁶.

¹⁹³ C. Rovane (1990), p.357.

¹⁹⁴ C. Rovane (1990), p.374-380. « In order to put the prospect of branching in the most favorable light, we ought to imagine a case where the subject actually has an investment in branching. We must suppose that from the dawning of self-consciousness – from the earliest time of the subject's reflective life – it has been understood what the future holds. In this scenario a subject's projects would have to be based on the prospect of branching, and it stands to reason that some of her projects would be undermined if she failed, in the future, to branch as expected." C. Rovane (1990), p.380.

¹⁹⁵ J'ai modifié les exemples. C. Rovane (1990), p.382-383.

¹⁹⁶ M. Siderits (2003), p.26-27, p.65.

Ces alternatives admettent tous la présence de sujets aux expériences et d'agents aux actions¹⁹⁷. Est-il possible de tenir compte de la dimension active, holistique et durable des expériences conscientes en les intégrant simplement dans une séquence d'événements mentaux désignés par « ceci » comme le suggère D. Parfit ? Je n'ai pas trouvé d'argument en ce sens dans la littérature entourant sa réflexion, mais une remarque de D. W. Shoemaker peut servir de piste à cette fin. C. Korsgaard reproche à D. Parfit de ne pas tenir compte d'une dimension autoritaire du tout sur les parties. Cependant, chez lui, des changements peuvent être connectés ou non avec la séquence à l'intérieur de laquelle ils s'inscrivent et cela a une fonction autoritaire¹⁹⁸. Dans ce cas, lorsque des contradictions surviennent à l'intérieur de « ceci », la prise de décision aura lieu en fonction de la présence ou de l'absence d'une connectivité avec cet ensemble. « Ceci » n'est donc pas contraint d'être conçu comme un agent, mais peut désigner une séquence d'événements mentaux *connectés*, cela rendant compte de la dimension autoritaire en d'autres termes.

3.4- Synthèse

Afin d'éviter un problème de circularité, Parfit a proposé des descriptions de processus conscients qui ne se réfèrent pas à la présence d'un moi identique à travers le temps. Il met en scène des scénarios où des relations similaires à celles qui unissent nos vies existent entre plusieurs personnes ou sont inclus dans des séquences désignées par « ceci ». Cette démarche traite l'esprit comme quelque chose qui se découpe en morceau isolable de l'ensemble, qui se présente sous un mode passif et dont l'unité est plus brève que la vie. Ces trois aspects de sa réflexion lui ont valu la critique selon laquelle les expériences conscientes dépendent de leur intégration dans l'histoire d'une vie sous le contrôle actif d'un agent pour avoir une signification. Il

¹⁹⁷ C. Rovane (1990), p.359 ne prétend pas défendre intégralement le réductionnisme de Parfit, car elle maintient la nécessaire référence à un sujet, mais elle argumente comme lui que la nature de ce sujet est dérivée de processus mentaux préalablement compris.

¹⁹⁸ D.W. Shoemaker (1996), p.327-328. La connectivité renvoie à l'intégration des expériences par des liens de mémoire, d'intentions, de traits de caractères et de croyances.

est nécessaire de se référer à l'identité d'un moi à travers le temps pour les comprendre. Trois répliques entourant ces problèmes ont été avancé à la défense de Parfit. Premièrement, des pathologies combinés au fait que le cerveau fonctionne par des systèmes séparés rendent concevable l'idée d'un processus mental similaire à la mémoire dont seulement l'identification au passé est absent. Deuxièmement, unifier l'ensemble de notre vie sous une narration n'est ni nécessaire, ni suffisant pour être le même à travers le temps. Troisièmement, des pronoms démonstratifs légèrement différents du « je » peuvent rendre compte de la dimension active et holistique de ces processus mentaux.

Dans ce qui suit, je clarifierai la discussion en évaluant sa forme, puis en réfléchissant à la présence de croyances métaphysiques non reconnues qui l'obscurcissent. Premièrement, l'usage de fictions présumant de certains faits scientifiques douteux contribue à alimenter ces objections. L'appel à la neurologie et à des cas pathologiques permet d'indiquer comment certaines méthodes plus fiables pourraient tester ces hypothèses. Cependant, avant même d'avoir des informations à cet égard, des faits plus familiers donnent du poids à la position de D. Parfit. Deuxièmement, la multiplication des contextes de signification et la reformulation d'une seule et même histoire de plusieurs façons sont des phénomènes assez banaux. Traiter la référence à un « moi » comme une exception à cette règle indique la présence d'une croyance métaphysique sous-jacente.

3.4.1- Observations courantes

Une difficulté posée par la réflexion de Parfit, c'est qu'elle s'appuie sur des fictions plutôt que des faits scientifiques. Il est improbable que le cerveau se découpe morceau par morceau, que la mémoire soit localisable, copiable et transplantable, puis que deux hémisphères équipotents existent. Les causes physiques de nos processus mentaux peuvent être si fortement intégrées les unes aux autres qu'il soit

impossible d'en transférer des parties à un autre système. La dépendance de la mémoire ou de la poursuite d'intention envers l'identité d'une chose à travers le temps est peut-être ainsi un fait.

Les données empiriques invoquées par G. Northoff, S. Klein, S. Nichols et R. Martin apportent une contribution faible à l'argumentaire de D. Parfit, car ils s'appuient sur quelques cas qui prêtent à interprétation. Il ne va pas de soi que la femme souffrant de schizophrénie avait un état mental similaire à la mémoire lorsqu'elle avait de l'information à propos d'une proche du tsar. Le témoignage d'un pur sentiment de dissociation devant un processus mental étant similaire est peut-être incohérent. Il nous faudrait plus d'observations sur des phénomènes similaires pour conclure à un système d'identification indépendant du reste de la mémoire. De plus, recouvrer la mémoire morceau par morceau ne montre pas que des états mentaux aussi riches que la mémoire normale peuvent se passer d'un contexte.

Ces réflexions sont intéressantes parce qu'elles indiquent des pistes pour confirmer les affirmations de D. Parfit. Si l'identification est un système neurologique distinct de ceux qui sont impliqués dans la mémoire et la poursuite d'intentions, alors cela constitue un argument en faveur de D. Parfit. En effet, la référence à un moi identique à travers le temps est alors traitée comme un phénomène neurologique dont nous pouvons observer l'absence en présence de tout le reste. De plus, la capacité d'encoder et de récupérer l'information à propos de la vie des autres via les mécanismes de la mémoire autobiographique peut également s'observer.

Cependant, avant même de le faire, des observations courantes plus facilement accessibles donnent du poids à l'approche de D. Parfit.

Premièrement, les changements de contextes narratifs sont constants à l'intérieur même d'une vie et cela suggère que la distance qui sépare plusieurs personnes sépare aussi les divers instants d'une même personne. Ce phénomène s'observe en relisant

un journal intime dix ans après l'avoir écrit. L'information fournie par la mémoire à propos du contenu qui y est décrit est pauvre en comparaison, voir déformée ou idéalisée. Nous n'accédons pas ainsi à la richesse des associations qui accompagnaient ces expériences même si, en réalité, ce sont les nôtres. C'est d'ailleurs le journal et non la mémoire qui aura intuitivement autorité. La voie d'accès la plus fiable au contexte lointain de notre propre expérience passée est en ce sens publique. Un individu qui cultive un sens fort de sa propre identité à travers le temps gagne à documenter sa vie d'albums photos, de journaux et de vidéos. De plus, discuter avec nos amis d'expériences communes corrige les failles de notre mémoire. Or, ces outils sont les mêmes que ceux qui nous relient au contexte des expériences d'autrui.

De ce point de vue, il est facile d'imaginer que la relation de Jane au voyage de Paul à Venise avancée par Parfit soit similaire à celle qui nous unit à nous-mêmes dix ans plus tôt. Dans les deux cas, le contenu est sorti de son contexte et appauvri. Après un certain temps, les sources d'informations publiques entretiennent une relation plus directe au contexte narratif original de l'expérience que les processus internes. Imaginer qu'une autre personne puisse quasiment mémoriser mes expériences est aussi facile à faire que d'imaginer un de nos proches en train d'en lire un récit très détaillé, puis ensuite s'en souvenir vivement, spontanément et ainsi de suite.

De plus, *l'unité de signification* de la vie est largement une illusion nourrie par des personnes qui ont une vision idéalisée d'eux-mêmes. La plupart du temps, nos vies ne concordent pas avec les fantaisies projectives ou rétrospectives que nous nous en faisons. Nous insufflons vaguement une direction abstraite au niveau de la carrière, du montant épargné pour la retraite ou de la forme générale de notre famille, mais toute la richesse des expériences qui en découlent n'est que très rarement l'objet d'une anticipation réussie. De plus, pour beaucoup de gens, aucune de ces projections ne tiennent la route. Nous n'en retenons ensuite que des grands traits

souvent peu représentatifs. L'idée que l'écrasante majorité de la saveur de la vie arrive tout simplement sans que nous ayons une prise là-dessus constitue en ce sens probablement une vision plus réaliste de la vie¹⁹⁹.

Deuxièmement, la phénoménologie de l'agir invoquée en support à la thèse que le point de vue de l'expérience consciente est celui d'un agent est plus simple que ces critiques le suggèrent. Lorsque j'agis, il n'est pas faux de dire qu'une action a lieu sans référence à un agent. Je ne me considère pas comme un agent pendant que j'agis. L'action a lieu d'une certaine manière et c'est lorsque je théorise après coup sur ce qui vient de se passer que j'y introduis ma présence. La distinction entre un mode passif et actif est trompeuse, car le mode passif suppose que j'observe des événements mentaux se produire devant moi. Bien sûr, je n'observe pas de cette manière mes pensées et mes actions défiler. Cependant, je n'accompagne pas non plus mon action d'une représentation de moi en train de la poser²⁰⁰.

De ce point de vue, la dimension active ou autoritaire consiste dans un flux d'actions automatiques, connectées et continues qui n'est pas médiatisé par une commande venant du monde extérieur. Mes pensées et mes actions s'enchaînent spontanément sans que je me représente une force intrusive qui vient l'obstruer. Lorsque j'ai à prendre une décision entre des options contradictoires après réflexion, le choix se fait à un moment donné. Cependant, je n'ai pas besoin d'introduire l'idée d'un moi qui donne une poussée à ce choix ou qui l'observe. En ce sens, décrire un changement qui est connecté à « ceci » décrit assez bien la réalité de l'action.

La fission peut illustrer cette réflexion. Une personne qui n'a jamais philosophé sur le problème de l'identité du moi à travers le temps risque d'envisager la fission sans même se poser de question. Elle imaginerait deux séries d'actions concurrentes l'une

¹⁹⁹ Cette réflexion va dans le sens de celle de D.W Shoemaker mentionné précédemment.

²⁰⁰ J.P. Sartres (1992), ainsi que M. Belzer (2005) défend une opinion similaire. Dans le 5^{ème} chapitre, j'analyserai plus en détail l'argumentation de ce dernier.

après l'autre. Son flux de conscience se diviserait en deux, puis ni l'être qui précède, ni les deux êtres qui suivent n'y verrait aucun problème. Ceux-ci agiraient conformément aux intentions qui les ont précédées. L'usage correct des pronoms pourrait leur apparaître comme un débat théorique sans importance. De nombreuses fictions cinématographiques impliquent ce genre de situation et les auditeurs ne sont pas portés à y voir de problème²⁰¹. Le fait que nous n'ayons pas à stipuler notre présence en tant qu'agent et que nos actions ont simplement juste à se produire explique pourquoi un tel scénario ne poserait aucun problème pour beaucoup de gens. Ils n'auraient pas à penser à l'identité des êtres concernés pour qu'une telle séquence d'actions se divise en deux²⁰².

3.4.2- Les intuitions sous-jacentes au postulat d'unité de signification

D'où vient cette croyance en une unité narrative si forte que toutes nos expériences conscientes doivent s'y référer pour avoir un sens ? Ceux qui s'en servent comme argument la postulent comme si elle allait de soi. Pourtant, les arguments précédents montrent que ce n'est pas le cas. Comment expliquer une telle erreur manifeste ? Il faut se rappeler que l'approche de D. Parfit ne vise pas à rendre compte de la perspective du moi telle qu'elle se présente à nous, car il considère qu'elle contient une illusion. C'est la croyance en l'existence d'une réalité au-delà du corps et des processus mentaux à laquelle le mot « moi » renvoie. Or, la spontanéité avec laquelle des philosophes affirment la présence d'une signification d'ensemble à la vie vers lequel tout le reste converge peut indiquer la présence de cette croyance implicite, bien que ceux-ci en nient la présence.

Comment justifier l'attribution d'une illusion métaphysique implicite à un auteur ? Si une personne change radicalement d'attitude relativement à une distinction mineure,

²⁰¹ C'est fréquent dans tous les récits de voyage dans le temps où les individus se fissionnent en plusieurs branches d'univers alternatifs. Star Trek contient de nombreux cas du genre.

²⁰² M. Belzer (2005), défend une thèse similaire.

cela indique vraisemblablement qu'elle lui attribue des propriétés magiques mystérieuses. Par exemple, si je prétends que mercredi n'est rien de plus qu'une journée comme une autre, mais que je ressens beaucoup d'anxiété à prendre des décisions financières ce jour-là plutôt qu'une autre, vous seriez en droit de soupçonner que j'entretiens des superstitions implicites à propos de cette journée. La même chose peut être dite du « moi ».

Lorsqu'il s'agit de raconter une histoire à la troisième personne ou à la première personne du pluriel, la possibilité de modifier légèrement les mots pour en parler ne pose aucune difficulté. Suis-je canadien ou québécois ? Est-ce que les décisions de l'assemblée nationale sont « nos » décisions ou celle d'un sous-groupe auquel je ne m'identifie pas ? Il est facile de réduire de tels débats à des attitudes plutôt qu'à des faits. Pourtant, lorsque le récit est fait par le « je », les auteurs qui s'appuient sur le contexte de narration pour critiquer Parfit supposent que nous ne pouvons pas jongler aussi librement avec les pronoms et les entités impliquées. La discussion précédente autour du « je* », de « ceci », d'une forme plurielle ou de la référence à un micro-moi bref illustre cette réticence. Or, ce changement draconien d'attitude envers la flexibilité des termes de la narration est étrange. Il semble que le moi agit comme une sorte de « abracadabra » narratif, excluant mystérieusement toutes les alternatives et suggérant qu'une modification verbale mineure sépare le réel en deux dimensions radicalement opposées.

Dans le même ordre d'idées, chacune des fois où nous lisons le même texte, nous y voyons une signification différente. L'objection du contexte narratif assume automatiquement que, lorsque ce texte est notre vie, ce n'est plus le cas. Il y a une seule signification et nous nous efforçons sans arrêt d'adapter nos actions en conséquence. Comme je l'ai fait remarquer plus tôt, cette thèse ne coïncide pas avec la réalité observable. Ici encore, le changement radical d'attitude suggère que le

« je » fonctionne comme un « abracadabra » éblouissant qui unifie par magie un ensemble dont l'apparence est manifestement multiple.

L'impression cartésienne d'avoir une autorité privilégiée sur la signification de nos expériences joue apparemment un rôle. Cette idée pose que le « moi » est muni d'une propriété « spéciale » inanalysable que n'ont pas les autres entités. Il faut être un moi pour en comprendre « l'effet que cela fait ». Il est raisonnable de penser que c'est une telle croyance non réductionniste qui nous pousse à proclamer spontanément l'unité narrative de la vie.

Il est révélateur qu'en dépit de la capacité de poursuivre l'ensemble du contexte narratif de notre vie en créant une copie parfaite de nous, de nombreux auteurs ne sentent pas que le résultat préserverait leur existence. Ils peuvent sentir qu'en réalité, le flux de la conscience cesserait. Or, c'est cette intuition qui n'est pas analysable et qui correspond à l'idée d'un moi non réductible.

L'argument de la fission peut être compris sous un autre angle. Ce qui semble bizarre dans ce cas, c'est de ne plus être autorisé à dire qu'il s'agit de mon passé, car quelqu'un d'autre me fait compétition pour dire la même chose. L'apparition de ce double fait alors que tous les états mentaux que j'ai actuellement me relient au passé d'une autre personne. J'ai pourtant beaucoup de difficultés à croire que l'étiquette « autre personne » accolé aux expériences que je me représente ouvre alors un fossé radical entre moi et lui qui n'existe pas entre moi et moi. La voie d'accès reste la même, mais elle est nommée différemment. Il est alors improbable que d'être relié à moi ou à un autre altère radicalement le contexte de significations qui nous unit.

Il est possible d'arriver à ce détour grâce à une observation courante. Reprenons mon exemple du journal intime écrit il y a dix ans. Lorsque je constate la distance qui me sépare aujourd'hui du contexte narratif qui l'entourait, il m'est difficile de prétendre que je possède une voie d'accès spéciale via la mémoire à la signification de ces

expériences. À ce moment-là, c'est la lecture qui m'en donne une vue plus fidèle. Ma mémoire apparaît comme une source d'information moins fiable. Pourtant ma voie d'accès la plus directe à ce moment-là est une voie à laquelle les autres ont accès : la lecture. Ce que nous nous représentons de cette manière peut être similaire. Dire que les autres mémorisent quasiment mon expérience est alors banal.

3.5- Résumé

D. Parfit affirme que notre existence consiste en l'existence d'un corps et d'un ensemble d'événements mentaux et physiques inter-reliés et non en une chose de plus. Afin de surmonter un problème de circularité, il décrit l'unité de la vie sans faire référence à un moi identique à travers le temps. Les arguments qu'il emploie pour y parvenir morcellent l'expérience consciente, puis en fait quelque chose de passif et de bref.

Ces aspects lui ont valu des critiques selon lesquelles il déforme considérablement leur signification, négligeant leur dimension holistique, active et durable. M. Schechtman lui reproche de traiter le moi comme un objet et non comme un sujet, ainsi que de séparer le contenu de la mémoire du contexte qui l'accompagne. Ce dernier est largement défini par la présence de grands projets unificateurs à l'intérieur desquels nous nous imaginons jouir de leur réalisation. A. Rudd reproche au réductionnisme de négliger la présence de patterns stables orientés dans une direction, accompagnés d'une quête de cohérence marquant une personnalité. T. Burge relie la présence du moi dans la mémoire à une capacité d'actualiser des projets durables. C. Korsgaard voit notre présence dans la nécessité de résoudre avec autorité des conflits afin de mener à terme des actions durables.

Plusieurs réponses à ces objections ont été avancées. Le cerveau traite les informations au moyen de systèmes séparés. Des cas cliniques suggèrent que l'identification peut fonctionner indépendamment de la mémoire autobiographique ou

morceau par morceau. D'autre part, l'intégration des expériences conscientes dans un récit durable n'est pas nécessaire pour exister, car de nombreux individus ne le font pas. De plus, elle n'explique pas la réticence à envisager notre survivance dans des scénarios comme la réplication. Finalement, nous pouvons introduire plusieurs pronoms alternatifs qui préservent la dimension holistique, active et durable de l'expérience consciente sans se référer au moi.

J'en ai conclu d'abord que les arguments empiriques sur le fonctionnement du cerveau peuvent fournir des pistes de confirmation plus crédibles que l'usage d'expériences de pensée, mais que l'usage de moyens publics pour préserver l'unité de signification à travers de longue période de temps et la simplicité de la phénoménologie de l'agir fournissent des raisons plus facilement accessibles en faveur de D. Parfit. Elles permettent de nier la différence radicale entre les changements de contexte narratif intra-personnel et interpersonnel, puis de nier la nécessaire référence à soi pour qu'une action ait lieu. Une manière d'expliquer la résistance à de telles conclusions, c'est de supposer une intuition non réductionniste.

CHAPITRE IV

UNITÉ DE L'EXPÉRIENCE EN UN INSTANT

Les trois précédents chapitres portaient chacun sur une manière d'être unifiée à soi-même à travers le temps. Dans ce chapitre-ci, nous réfléchirons plutôt sur une façon de l'être en un seul instant : l'unité phénoménale.

L'unité phénoménale correspond à ce que T. Nagel qualifie « d'effet que cela fait »²⁰³. Il y a un « effet que cela fait » pour une chauve-souris de se déplacer dans l'espace au moyen de l'échololocation, mais il n'y en a pas pour une roche de se mouvoir. De la même manière, il y a un effet que cela fait *pour moi* de voir une plage et d'entendre le bruit des vagues en même temps. Par contre, si je les vois en étant sourd à côté d'un aveugle qui les entend, il n'y aura pas un effet que cela fait *pour nous* de les voir et de les entendre *ensemble*²⁰⁴. Notre « moi » est donc la cible de cet « effet que cela fait ». Cela le distingue du « nous ». L'ingrédient fondamental qui sépare les êtres conscients est ainsi cette union des expériences²⁰⁵.

Certains y voient le signe qu'une chose inanalysable se trouve dans notre expérience au-delà de nos processus mentaux²⁰⁶. D. Parfit a répliqué que cette unité ne correspond pas nécessairement à la présence d'un moi et peut être décrite sans s'y référer²⁰⁷. Son argumentation combine à la fois des considérations empiriques autour

²⁰³ T. Nagel (1981).

²⁰⁴ J'ai modifié les exemples de T. Nagel.

²⁰⁵ Pour une description de ce type d'unité phénoménale et de sa relation au moi, voir T. Bayne et D. Chalmers (2003)

²⁰⁶ R. Swinburne (1984), p.42-48, J. Ganeri (2000) et A. Gilead (2008) défendent une telle argumentation contre le réductionnisme de D. Parfit. B. Dainton (2004, 2008), D. Zahavi (2000) et W. Fasching (2009) ne se réfèrent pas à une *chose* « supplémentaire » au corps, mais la récupèrent pour montrer que le corps possède une *dimension* supplémentaire à ce que D. Parfit reconnaît dans son ontologie.

²⁰⁷ D. Parfit (1984), p.216. Cette unité est réelle selon lui, mais elle n'a pas besoin du moi pour être décrite : « When Nagel asked his famous question « what is it like to be a bat ? » he meant « what it is like *for the bat* ? » To understand the qualitative character of experience, it may be claimed, we must have the concept of the subject *for* whom these experiences have this character. Since my beings have

de cas cliniques et l'imagerie de la science-fiction. Les critiques qui lui sont adressées portent sur l'interprétation de ces cas réels ou sur l'intelligibilité de son imaginaire. La plupart d'entre elles ne font pas intervenir une chose de plus que le corps, le cerveau et l'expérience consciente, mais présentent cet aspect comme étant indissociable de la référence à soi. Je présenterai d'abord l'argumentation de Parfit, pour discuter ensuite de pour et de contre. Je diviserai cette discussion en deux aspects, empiriques et conceptuels. J'argumenterai que des cas cliniques supportent la position de Parfit et que l'utilisation fréquente d'un mode public d'autoréférence rend possible son imaginaire. En fait, notre inconfort face à ces conclusions repose plus vraisemblablement sur des croyances mystérieuses (non réductionnistes) à propos du rôle causal ou de la présence de soi dans l'expérience consciente.

no concept of a subject, they could not understand what experiences were *like*. This objection seem to me mistaken. Though these beings do not think of experiences (...) as had by subjects, that would not, I believe, make them unable to understand the qualitative character of these conscious states and events." D. Parfit (1999), p.256-257.

4.1- L'argumentation de Parfit

D. Parfit a proposé une fiction intitulée « l'examen de physique » afin de montrer comment l'unité phénoménale des expériences peut être décrite sans se référer à soi²⁰⁸. Nous devons nous imaginer en train d'effectuer deux séries de calcul contenant séparément une unité consciente. Il y a alors un « effet que cela fait » dans l'hémisphère droit de calculer et un « effet que cela fait » dans l'hémisphère gauche de calculer, mais il n'y a pas « d'effet que cela fait » pour moi de les calculer en même temps²⁰⁹. J'ai alors deux flux de conscience simultanément. Donc, l'unité de chacun d'entre eux ne peut pas être décrite au moyen d'une référence au fait que je les vive, car cela les confondrait en un seul²¹⁰. Dans cette situation, le moi échoue à les unir phénoménologiquement. Donc, pour décrire l'unité dans chacun des flux, nous devons pointer seulement l'état des expériences. Leur unité se décrit alors mieux comme un état de conscience singulier de plusieurs expériences co-conscientes. « It is simply a fact that several experiences are co-conscious, or be the objects of a single state of consciousness.²¹¹ »

Il y a deux significations concurrentes à la notion de « moi » dans cette fiction de D. Parfit. « Je » peut coïncider avec le réseau de pensées et d'expériences qui s'étend sur de longues années. Dans ce cas, « je » ne renvoie pas à une chose unifiée par un flux de conscience *en un instant*, mais à une chose unifiée consciemment *dans le*

²⁰⁸ D. Parfit (1984), p.246-247.

²⁰⁹ D. Parfit ne parle pas d'unité phénoménale, mais comprend le « flux de conscience » comme une manière d'être informé de nos actions de l'intérieur, cherchant à imaginer « l'effet que cela fait » de se diviser et de se réunir : « Consider my experiences in my « right-handed » stream. I remember deciding that I would use my right hand to do the longer calculation. (...) In working at this calculation I can see, from the movement of my left hand, that I am also working at the other. But I am not aware of working at the other. I might, in my right-handed stream, wonder how, in my left-handed stream, I am getting on. I could look and see. This would be like looking to see how well my neighbour is doing, at the next desk. (...) I suggest we can also imagine *what it would be like* to divide and reunite our minds. » D. Parfit (1984), p.247.

²¹⁰ « We need to explain the unity of consciousness within each of my two streams of consciousness, or in each of my divided mind. We cannot explain these two unities by claiming that all of those experiences are being had by me at this time. This make two unities one. » D. Parfit (1984), p.249.

²¹¹ D. Parfit (1984), p.250.

temps, mais *temporairement désunie* pendant 10 minutes. C'est dans ce cadre que nous ne devons pas décrire l'unité de la conscience en référant à mon état, mais à l'état des expériences. Cependant, « je » peut aussi vouloir dire « la chose qui est phénoménalement unifiée ». Dans ce cas, il y a deux « je » à l'intérieur d'un corps. Leur existence dure 10 minutes²¹².

D. Parfit reconnaît ces deux possibilités et y voit un cas d'indétermination de l'identité du moi *en un seul instant*. Au même titre que l'opposition entre une conception du moi centrée sur une chose physique ou sur la continuité d'une psychologie ne repose pas sur un fait qui transcende ces composantes, l'opposition entre ces deux usages du « je » ne porte pas sur la réalité, mais sur un choix de descriptions. Il estime cependant plus simple de considérer, dans son exemple, un seul être conscient temporairement divisé que de faire apparaître deux nouvelles entités d'une durée de 10 minutes²¹³.

Donc, sous la description la plus crédible, l'unité des deux séries d'expériences ne peut pas être décrite comme l'état d'une chose consciente ; moi. Pendant 10 minutes, « je » est aussi impotent à unifier diverses expériences sous une conscience qu'une nation, un couple, une amitié ou une famille puisse le faire. « Je » existe sous la forme d'une collectivité composée de deux consciences séparées l'une de l'autre. Chacune d'entre elles ne reçoit pas de moi leur unité phénoménale.

C'est de cette manière que D. Parfit s'y prend pour décrire l'unité phénoménale sans se référer à soi et montrer qu'elle n'indique pas la présence d'une chose de plus au-delà du corps et d'un ensemble de processus mentaux inter-reliés²¹⁴.

²¹² Pour une telle distinction, voir T. Alter (2010), p.29. Je montrerai plus loin comment T. Bayne (2010) s'appuie sur cette distinction pour critiquer une approche comme celle de D. Parfit.

²¹³ D. Parfit (1984), p.248.

²¹⁴ « On the Cartesian View, a particular mental event occurs within a particular life solely in virtue of its ascription to a particular Ego. We can deny that the topography of "Mental Space" is given by the existence of such persisting Egos. We can claim that a particular mental event occurs within life in

4.2- Possibilités empiriques d'une division de l'expérience consciente

4.2.1- Les cerveaux divisés

Est-ce que l'argumentation de D. Parfit a un fondement empirique ? Il cherche à montrer que oui au moyen des cas de cerveaux divisés.

La possibilité de l'examen de physique pouvant être contestée, il l'appuie sur des cas de césure du corps calleux reliant les deux hémisphères. Les patients manifestaient ensuite des signes de division de l'activité consciente en situation de tests expérimentaux. Par exemple, le patient devant identifier ce qu'il voyait niait verbalement voir quelque chose, mais retirait au même moment avec sa main une pipe d'une série d'objets de manière à indiquer qu'il la voyait. L'explication fournie, c'est que le contrôle de la parole et d'un champ de la vision formait un flux de conscience indépendant du contrôle moteur de cette main et du champ de vision incluant cette pipe²¹⁵. Il s'agirait donc d'un véritable cas de division de la conscience phénoménale démontrant la possibilité du scénario de D. Parfit²¹⁶.

Selon T. Bayne, cette division n'est pas simultanée²¹⁷. G. Gillett objecte que cette division est trop rare et pas assez complexe pour justifier les conclusions qu'en tire D. Parfit²¹⁸. M. Tye, quant à lui, croit que les comportements apparemment incohérents des personnes dans ces situations s'expliquent plus facilement par la division de l'expérience consciente²¹⁹. M.D. Reid suggère que les souvenirs rétroactifs de personnes aux prises avec un problème de stress post-traumatique possèdent les

virtue of its relations to many other mental and physical events which, by being interrelated, constitute this life". D. Parfit (1984), p.252.

²¹⁵ D. Parfit (1984), p.246. Pour une description plus détaillée de ce cas, voir T. Nagel (1971).

²¹⁶ « It might be objected that my description ignores « the necessary unity of consciousness ». But I have not ignored this alleged necessity. I have denied it. What is a fact must be possible. And it is a fact that people with disconnected hemispheres have two separate streams of consciousness." D. Parfit (1984), p.247.

²¹⁷ T. Bayne (2010).

²¹⁸ G. Gillett (1986).

²¹⁹ M. Tye (2003).

caractéristiques centrales de l'argument de l'examen de physique²²⁰. À la lumière de cette discussion, je soutiendrai que l'utilisation de ces cas est justifiée et que le comportement interprétatif impliqué dans cette situation peut nous informer d'une croyance métaphysique illusoire implicite.

4.2.2- Critiques de leur utilisation

T. Bayne conteste l'interprétation que D. Parfit fait des cerveaux divisés en mentionnant qu'une étude subséquente montre l'absence d'activités conscientes dans un hémisphère lorsque l'autre en supporte une. La rapidité des changements d'activités d'un hémisphère à l'autre donne l'illusion qu'il y a deux activités simultanées. Il explique donc ce phénomène par une alternance rapide du flux de conscience d'un hémisphère à l'autre. Cette théorie justifie aussi pourquoi le patient ne manifeste ces signes de division qu'en situation expérimentale²²¹.

G. Gillett conteste également l'interprétation de D. Parfit. Ce dernier en simplifie les résultats. Les indices sonores ne mènent pas à une division. Celle-ci n'a lieu qu'en situation expérimentale, donc il n'y a probablement pas deux flux de conscience autonomes l'un de l'autre. De plus, les patients s'attribuent à eux les erreurs et non à des champs phénoménaux séparés. Ils combattent activement pour rétablir l'unité. Finalement, les conditions d'une pensée complexe ne sont présentes que dans un seul des hémisphères. Il est donc plus probable que les incohérences ne soient pas conscientes dans l'autre²²².

4.2.3- Support à leur utilisation

En réponse à T. Bayne, M. Tye soutient que la meilleure manière d'expliquer la division du comportement des patients est la division de leur phénoménologie et non

²²⁰ M. D. Reid (2005).

²²¹ T. Bayne (2010), chapitre 9.

²²² G. Gillett (1986), p.224-227

l'alternance inter-hémisphérique de la conscience. Premièrement, les tâches qu'ils performent séparément sont isolément suffisantes pour attribuer la conscience à des enfants en phase pré-linguistique. Deuxièmement, il s'appuie sur les expériences de vision aveugle. Un sujet devine la présence d'un verre lorsque nous le questionnons, mais il ne le saisit pas lorsqu'il a soif. Dans ce cas-là, la meilleure explication de l'incohérence apparente est l'absence d'une expérience consciente du verre. Les réponses contradictoires venant d'un patient au cerveau divisé sont une extension de ce cas. De plus, M. Tye estime que de s'imaginer l'effet que cela fait d'avoir une expérience d'un côté, une expérience de l'autre et non dans les deux, ne lui pose pas de problème. Contrairement à D. Parfit, il défend cependant que cette division phénoménale n'est que partielle²²³.

M.D. Reid a suggéré qu'un processus similaire explique un phénomène relatif au stress post-traumatique. Lors de l'événement traumatisant, la personne affirme ne pas éprouver de sensations, mais dit ensuite s'en souvenir comme ayant eu une charge émotive très forte. Il suggère que les deux hémisphères, au moment de l'événement, ont de façon stratégique séparé l'expérience émotive de l'expérience cognitive en brisant le lien conscient entre les deux hémisphères. Ils les ont réunies en une seule conscience après-coup, expliquant le « flashback » différent du rapport d'origine²²⁴. Dans ce cas, le cerveau fait la même chose que ce que D. Parfit décrit, sauf qu'il ne divise pas deux expériences de calcul, mais une expérience émotive et cognitive²²⁵. M. D. Reid invoque en preuve des expériences sur la latéralité. Une personne sous sédation d'un hémisphère donne des contre-rendus différents du passé si c'est celui de gauche ou de droite qui est ainsi, suggérant qu'elles ont été vécues différemment. Le contrôle prédominant d'un des hémisphères sur le langage occulte ce fait²²⁶.

²²³ M. Tye (2003), p.117-132.

²²⁴ M. D. Reid (2005), p.693-694.

²²⁵ M. D. Reid (2005), p.689-690.

²²⁶ M. D. Reid (2005), p.686-689. « Our central nervous systems endow us with a seemingly paradoxical ability: to experience "without experiencing". That is, we can "feel significance", which

4.2.4- Synthèse de la discussion

L'imaginaire proposé par D. Parfit est-il confirmé par des cas réels ?

4.2.4.1- Ramener les objections au but visé

G. Gilett mentionne que les patients ne manifestent pas une division de l'activité consciente comparable à ce que D. Parfit nous demande d'imaginer. Leurs deux activités conscientes séparées ne se réfèrent pas à elles-mêmes et ne maintiennent cette division que brièvement. En ce sens, il n'y a aucune raison empirique de penser que le cerveau a la capacité de subir le genre d'opérations imaginées par D. Parfit.

Est-ce que cette objection lui est fatale ? Cela dépend de son objectif. Comme de nombreux arguments employés par les philosophes de l'identité personnelle, ils visent à montrer, au moyen de l'imagination, comment deux phénomènes constamment associés peuvent être conçus séparément²²⁷. Dans ce cas-ci, c'est la description de l'unité phénoménale des expériences et l'attribution à un moi que Parfit vise à séparer. Elles sont habituellement si inter-reliées que cet exercice peut sembler inconcevable. La présentation des cas de cerveaux divisés sert alors à modifier cette attitude. En ce sens, même si une division de l'esprit en deux flux munis de toutes les caractéristiques d'une conscience individuelle est physiquement impossible, la division de la conscience phénoménale en un bref instant apporte un support à son raisonnement.

promotes social bonds and self-concern, while at the same time using self-identification mechanisms in way that favor the conscious contents of the left hemisphere. It is a mistake to infer that right hemisphere processes are non-conscious at these times of high significance, since evidence shows that the right hemisphere plays role in representing significance. (...) The right hemisphere may be no more "sub" or "un" conscious than the left hemisphere, but only dissociated from the left hemisphere, which makes self-judgements, has most self-structures, and generally provides clearest behavioral evidence of the content of consciousness". M. D. Reid (2005), p.689.

²²⁷ Pour une telle explication, voir D. Kolak (1993).

Les remarques de M. Tye supportent cette conclusion. Le fait que, de façon simultanée ou quasi-simultanée, une personne puisse affirmer verbalement ne pas voir de pipe et faire signe de la main en agitant une pipe constitue deux comportements possédants toutes les caractéristiques d'expériences conscientes. La contradiction indique apparemment l'absence d'une phénoménologie commune entre deux parties d'un être. Cette interprétation est intuitive. D. Parfit atteint donc son objectif, car nous réussissons plus facilement à séparer conceptuellement l'attribution à soi et l'unité phénoménale en un instant. Même si, après une analyse plus approfondie, cette réponse est inadéquate et qu'il n'y a aucune division de ce genre, nous sommes contraints d'admettre qu'elle est intelligible et c'est là son seul but. Les arguments de G. Gillet indiquent des raisons valables de douter de la présence d'une telle désunité, mais ils n'éliminent pas l'intelligibilité de l'hypothèse adverse.

La suggestion de M. D. Reid d'expliquer des réactions en situation de stress post-traumatique par division et réunion de la conscience renforce cette conclusion, car il s'agit alors d'une division stratégique plus complexe des tâches dont la rationalité est plus facilement attribuable à un système psychologique qui dure longtemps qu'à l'apparition de deux nouveaux êtres conscients éphémères. Effectivement, lors du traumatisme, c'est l'individu persistant dans le temps qui cherche à se protéger en divisant sa conscience et non deux micros sujets. C'est donc intégralement lui qui vit deux expériences conscientes distinctes et simultanées. Même si son explication est fausse, elle demeure intelligible et c'est tout ce dont a besoin D. Parfit.

4.2.4.2- Les hypothèses superflues visant à camoufler le malaise

La réponse de T. Bayne montre la présence d'une alternative à l'interprétation des cas de cerveaux divisés par division phénoménale de la conscience en un instant : une rapide alternance inter-hémisphérique.

Le problème, c'est que cette réponse est superflue. Il n'y a pas de différence significative entre une brève succession de deux expériences que nous n'unifions pas consciemment et une parfaite simultanéité de deux expériences que nous n'unifions pas consciemment. Dans les deux cas, il y a absence de l'unité de la conscience phénoménale là où nous nous attendons à en trouver habituellement une. Le patient lui-même ne témoigne pas d'une sensation de changer de champs de vision à la manière d'un conducteur qui change de véhicules. Dans les cas normaux, nous avons cette impression de vivre le passage d'une expérience consciente à l'autre²²⁸.

L'hypothèse de T. Bayne suppose erronément que la conscience est une chose dont le sujet observe intérieurement le *déplacement*. Or, les signes comportementaux de la personne au cerveau divisé suggèrent seulement une succession *inconsciente* de deux expériences conscientes²²⁹.

4.2.4.3- La signification des comportements interprétatifs

Lorsque D. Parfit utilise des fictions, c'est pour nous révéler une disposition à croire en l'existence d'une chose au-delà des composantes que nous mettons en opposition pour définir ce que nous sommes²³⁰. Ensuite, c'est pour contester que nos processus

²²⁸ B. Dainton (2004) analyse abondamment cette sensation, croyant y trouver l'essence de l'identité du moi à travers le temps.

²²⁹ T. Bayne (2010), p.219 envisage cette objection et y répond en affirmant que ce qui individualise le flux de conscience est l'absence d'interruption. Ce critère est inadéquat, car les données observables nous suggèrent précisément la présence d'une telle interruption.

²³⁰ « Some philosophers object that, since our concept of person rests on a scaffolding of facts, we should not expect this concept to apply in imagined cases where we think those facts away. I agree. But I believe that, for a different reason, it is worth considering such cases. We can use them to discover, not what the truth is, but what we believe. We might find that, when we consider science fiction cases, we simply shrug our shoulders. But that is not so. Many of us find that we have certain beliefs about what kind of fact personal identity is." D. Parfit (1995), p.15. "In our thoughts about our identity, we are prone to illusions. That is why the so-called "problem cases" seem to raise problem: why we find it hard to believe that, when we know the other facts, it is an empty or merely verbal question whether we shall exist." D. Parfit (1995), p.45.

mentaux nous en révèlent la présence²³¹. Son utilisation des cas de cerveaux divisés s'inscrit dans la même logique et doit être évaluée dans ce cadre.

Nos comportements face aux effets visibles d'une césure du corps calleux peuvent être révélateurs. La tendance marquée de plusieurs à chercher au-delà de la division apparente des explications, des interprétations ou des signes rétablissant la présence d'une unité phénoménale nous informe de quelque chose : s'ils ont devant eux des indices d'une désunion de la conscience de soi, ils ne vont pas la reconnaître comme possible en haussant les épaules, mais continuer de fouiller dans la situation pour la nier. Ils ont donc le réflexe non réductionniste de ne pas se contenter des composantes observables impliquées dans la situation pour tirer des conclusions, mais de chercher une chose au-delà.

La réponse de T. Bayne est intéressante à cet égard. Elle est invoquée pour préserver l'intuition d'une nécessaire association entre l'attribution à soi et l'unité phénoménale et n'a pas d'autre fonction explicative. La simple succession de deux expériences conscientes en présence de signes de désunité ne justifie pas cette addition théorique. Seulement son souci de préserver une intuition le fait. Il nie croire en l'existence d'une chose consciente transcendant la division des hémisphères et la brève succession des expériences²³². Pourtant, cette pulsion d'y chercher une explication *supplémentaire* devant une réalité qui n'en a pas besoin indique vraisemblablement qu'il possède à un niveau plus ou moins conscient cette croyance qu'il prétend ne pas avoir²³³.

²³¹ Voir chapitre 1.

²³² Il parle du moi en tant que "fiction virtuelle". T. Bayne (2010), p.269-294.

²³³ L'interprétation du comportement humain demande une certaine méfiance relativement aux énoncés qu'une personne exprime relativement à ses propres états mentaux. Fréquemment, nous questionnons la conscience qu'elle en a en constatant une dissonance entre ses dires et ce qu'elle nous montre par ses gestes. La même chose peut être dite du *comportement philosophique* d'un auteur. Chercher une explication sans rapport logique avec les faits présentés de manière à prouver la présence d'une nécessité m'apparaît être particulièrement suspect. Nous pouvons soupçonner raisonnablement que l'auteur s'en fait une idée dont il n'est pas entièrement conscient.

4.2.5- Résumé

En résumé, la présence d'un « effet que cela fait » pour nous de vivre une seule et même expérience composée de multiples facettes en un instant est intuitivement associée à notre présence. Elle nourrit l'intuition de notre irréductibilité. Afin de remettre cette idée en question, D. Parfit a argumenté que ce phénomène se décrit sans se référer à soi. Il imagine une situation où nous divisons notre conscience en deux afin d'effectuer deux séries de calculs séparément. Il appuie l'intelligibilité de ce scénario sur les cas de césure du corps calleux, qui avait comme conséquence une division mentale apparente. Selon lui, dans ces cas, il est préférable de décrire l'unité des expériences conscientes en pointant l'état des expériences et non soi-même afin d'éviter de confondre deux expériences distinctes en une.

Certains lui ont objecté que les cas de césures du corps calleux ne manifestent pas des signes de division assez complexes pour justifier ces conclusions. De plus, il n'y a pas de signe de division simultanée, mais en succession. Selon T. Bayne, il y a plutôt une alternance de l'activité consciente.

Dans la mesure où l'objectif de D. Parfit est de montrer comment deux phénomènes peuvent être conçus séparément, l'explication par division phénoménale est suffisamment crédible pour justifier sa conclusion. De plus, M. D. Reid a expliqué comment une personne aux prises avec un problème de stress post-traumatique manifeste des signes de division suffisamment complexes pour ressembler au scénario de D. Parfit, le rendant concevable. D'autre part, l'explication par alternance n'ajoute rien par rapport à la description de D. Parfit, car une brève succession de deux expériences ne se distingue pas d'une simultanéité.

Cette tendance à nier la possibilité d'une division phénoménale de la conscience en présence de signes contradictoires semble être mue par une intuition non-reconnue selon laquelle nous sommes une chose au-delà du corps et des processus mentaux

particuliers. C'est dans ce cadre que cette discussion prend sa pertinence et ce, indépendamment de ce en quoi consiste réellement l'effet d'une césure du corps calleux.

4.3- Possibilité conceptuelle d'une division de l'expérience consciente

4.3.1- Moi défini comme entité à double face, publique et privée

Dans la section précédente, nous avons évalué s'il y avait des raisons empiriques de croire que la division de l'expérience consciente en une seule personne était possible. Au-delà de cette difficulté, il se peut que ce scénario nous semble plausible seulement parce que nous ne précisons pas le sens des concepts. Un aspect ou un autre de sa description ne respecte peut-être pas l'usage de concepts clés entourant l'unité des expériences. Par exemple, certains croient qu'un moi est *par définition* une entité qui se réfère à elle-même via une dimension privée.

Un élément qui fait obstacle à nous imaginer divisant notre conscience en deux, c'est l'idée de nous observer de l'extérieur. En effet, « je » suis censé voir mon champ de vision perdre la moitié de son contenu, constater des mouvements de crayons, y voir une activité consciente de l'autre côté de mon cerveau, puis en conclure que « je » existe alors à deux endroits en même temps. Cela se produit sans que j'aie accès à l'effet que cela fait d'être moi de l'autre côté. « Je » suis alors informé de l'étendu de mon existence par une voie d'accès extérieure et non intérieure. En percevant des indices de calculs de l'autre côté, une partie de moi observe l'autre partie de moi comme elle observerait une autre personne. « Je » ne me sens pas directement être à deux endroits, mais « je » le déduit de mes cinq sens de chacun des côtés.

Je peux employer alors le « je » comme synonyme de « nous ». Ce à quoi renvoie « je » n'est alors pas moi, mais autre chose et je joue avec les mots²³⁴. « Je » est censé renvoyer à une entité capable de se percevoir de l'extérieur et de l'intérieur. C'est ce que J. L. Borges illustre dans son texte « Borgès et moi ». Il désignait par « Borgès » tous les aspects publics de son existence, comme être un écrivain ou marcher dans Buenos Aires. Ensuite, il distingue « lui », l'être associé à Borgès, mais qui ne se réduit pas aux prédicats observables qui le définissent²³⁵. T. Nagel exprime une idée similaire²³⁶. Je peux savoir qu'une chauve-souris s'oriente dans l'espace au moyen de l'écholocation en étudiant ses comportements, mais je ne peux pas avoir accès à l'effet que cela lui fait de se mouvoir de cette manière. Son « intérieur » m'est inaccessible. Le « je » de la chauve-souris est donc composé d'une face extérieure que je peux étudier et d'une face intérieure qui m'est voilée.

Le « je » que D. Parfit nous demande d'imaginer dans l'examen de physique est dépourvu de ce double aspect. Il est une entité purement publique. Il fonctionne comme un « nous ». Si je suis aveugle et entend les vagues et que tu es sourd, mais les vois, nous ne formons pas ensemble une entité munie du pouvoir de vivre intérieurement les vagues sous ces deux aspects. En ce sens, « nous » n'avons pas un intérieur en commun. Or, lorsque « je-1 » m' imagine observer au-delà de mon champ de vision « je-2 » faire un autre calcul, les deux « je » sont deux intériorités distinctes et « je-3 » censé être les deux en même temps est en fait un « nous » formé de « je-1 » et « je-2 » et, lui, il n'a pas plus d'intérieur que le nous composé d'un sourd et d'un aveugle en train de percevoir des vagues. D. Parfit échoue donc à imaginer un moi possédant deux flux de conscience, mais réussit seulement à nous

²³⁴ J. Ganeri (2000), p.654-655 reproche ainsi à D. Parfit d'éviter la référence à soi simplement en paraphrasant le moi dans d'autres mots, tout en préservant la même idée: « (...) if the only thing that makes such paraphrases succeed is that the aggregate is identified as the reference of a token of the first-person, then the paraphrase is not genuinely impersonal, for the concept of a person is being used to demarcate aggregates. »

²³⁵ J. L. Borgès (1981), p.19-20.

²³⁶ T. Nagel (1981), p.391-403.

faire imaginer un nous constitué de deux moi dans un corps et de lui substituer l'étiquette « je ».

Ces critiques lui sont adressées de trois manières.

D'une part, T. Bayne affirme que le moi se définit comme le centre d'une perspective phénoménale, c'est-à-dire l'intériorité décrit par J.L. Borgès et T. Nagel. Nous devrions donc conclure dans l'examen de physique qu'il y en a deux. Il refuse de considérer le système de processus mentaux qui s'étend sur plusieurs années comme un référent potentiel du « je » si ce système est dépourvu de cet aspect, le postulant comme une condition essentielle²³⁷.

D'autre part, T. Bayne, W. Fasching et J. Ganeri définissent le moi comme l'entité qui use en un instant d'un mode d'autoréférence par voie interne et directe²³⁸. En ayant accès à l'ensemble des aspects de notre expérience, nous pouvons nous référer automatiquement à nous-mêmes. Or, la manière dont « je-dans l'hémisphère gauche » renvoie à « je-dans l'hémisphère droit » use d'intermédiaires publics. Cette autoréférence diffère de ce qui spécifie le mode à la première personne du singulier²³⁹.

De plus, « l'état de conscience singulier » que D. Parfit décrit dans chacun des hémisphères possède la capacité de s'attribuer la multiplicité des aspects de

²³⁷ T. Bayne (2010), p.280: « What it would be like to enjoy the one stream of consciousness would be distinct from what it would be like to enjoy the other stream of consciousness. Or, to put the point more carefully, since we cannot think of self as having multiples perspectives, we should take there to be two perspectives in a disunified mind, each of which is had by a distinct self. A disunified mind would not itself have any conscious experience. »

²³⁸ J. Ganeri (2000). W. Fasching (2009), p.143-144. D. Zahavi (2000) apporte un argument similaire.

²³⁹ T. Bayne (2010), p.280. J. Ganeri (2000) le critique ainsi en s'appuyant sur l'idée que le moi tel que défini par D. Parfit implique qu'il construit son unité en un instant par de tels moyens externes, alors qu'il le fait de façon interne.

l'expérience sous le mode propre à la première personne du singulier, donc il a toutes les conditions pour être traité comme un être un moi à part entière²⁴⁰.

4.3.2- Autoréférence à la première personne occasionnellement publique et divisible

Les objections précédentes supposent que la relation entre moi et l'unité des expériences est une relation de nécessité conceptuelle. Par définition, je suis celui pour qui il y a un effet que cela fait de vivre une expérience et qui se réfère à lui-même en un instant sous ce mode d'appréhension. S'il n'y a pas un, mais deux effets, alors il y a deux êtres. Je suis aussi le propriétaire des états mentaux, c'est-à-dire la chose dont ils sont des états. Ainsi, s'il y a un « état de conscience singulier » de plusieurs expériences, cet état est un moi, car il en est le propriétaire. Ces difficultés conceptuelles posées par la fiction de D. Parfit se ramènent à la nature de l'autoréférence. Je suis censé pouvoir me représenter moi-même ici et maintenant de l'intérieur grâce à une expérience d'ensemble de toutes les parties qui me composent. Cette expérience étant absente dans l'examen de physique entre les expériences des deux hémisphères, il n'y aurait pas un moi, mais deux.

Le problème avec des objections de ce genre, c'est de prétendre trop rapidement connaître la signification ultime des termes employés, alors qu'il est encore possible au contradicteur d'apporter des précisions supplémentaires. Affirmer qu'il ne peut pas le faire, c'est équivalent à le déclarer perdant avant de le laisser répondre.

Or, D. Parfit n'affirme pas que je suis *toujours* une entité qui se représente sa propre existence de façon publique sous le modèle de l'examen de physique. Il affirme simplement que je pourrais *temporairement* me diviser, puis adopter, en un même instant, ce mode d'autoréférence pour unir les deux parties de moi-même. Cet énoncé est compatible avec le fait que, *en règle générale*, je me réfère à moi-même au présent via l'unité phénoménale de mes différentes parties. C'est pour cette raison

²⁴⁰ T. Bayne (2010), p.279. W. Fasching (2009), p.143.

qu'il avance la fiction d'une division *brève* de son activité consciente et non d'une division *permanente*²⁴¹. Cette distinction est suffisante pour l'amener à atteindre ses objectifs : montrer comment l'unité des expériences peut être comprise sans se référer à soi, que ce sont deux entités distinctes et donc qu'elle ne nécessite pas l'intervention d'une chose consciente simple et indivisible au-delà de l'expérience pour avoir lieu.

C. Rovane apporte des nuances relativement à l'autoréférence qui supportent D. Parfit²⁴². Dans son approche, « Je » désigne un engagement à réaliser une unité d'ensemble sur le long terme. Cela suppose de maintenir dans le temps un certain nombre de connaissances²⁴³. Le « je » se distingue actuellement des autres entités au « nous » par le degré d'intégration possible²⁴⁴, mais il est possible qu'une entité collective puissent se référer à elle-même au « je » de cette manière. C. Rovane n'attribue pas au « je » nécessairement une unité phénoménale²⁴⁵. Ainsi, une unité d'ensemble peut être réalisée en employant des moyens publics comme un cahier de bord au lieu d'employer un moyen privé comme la mémoire. Le fait que nous maintenions la référence à nous-mêmes à travers le temps aux moyens d'écrits est un exemple d'une telle autoréférence publique à la première personne du singulier²⁴⁶. Ensuite, il est conceptuellement possible de s'engager dans un projet cohérent qui

²⁴¹ « If we are not yet Reductionists, as I shall assume, we believe that it is a real question whether such cases involve more than a single person. Perhaps we can believe this in the actual cases, where the division is permanent. But this belief is hard to accept when we consider my imagined Physics exam. In this case there are two streams of consciousness for only ten minutes. And I later seem to remember doing both calculations that, during these ten minutes, my two hands could be seen to be writing out. Giving the brief and modest nature of this disunity, it is not plausible to claim that this case involves more than one person." D. Parfit (1984), p.248.

²⁴² C. Rovane (1990 et 1993)

²⁴³ C. Rovane (1990), p.357 et (1993), p.76..

²⁴⁴ C. Rovane, comme D. Parfit, parle de cette intégration en termes de connectivité psychologique.

²⁴⁵ C. Rovane (1990), p.363, note 8.

²⁴⁶ C. Rovane (1993), p. 96: " Many academics – including myself – who write books are in roughly the same situation as the human being just described, i.e, they must keep something like the log I mentioned in order to pursue their long-term activities. So I repeat my question. Should I be denied a first-person relation to past thoughts and actions simply because I must make a deliberate effort to retain my knowledge on them in written form ?"

implique de se diviser²⁴⁷. Pour préserver un mode d'autoréférence apparenté au « je », il suffit d'employer un pronom « je* » suivi d'une description²⁴⁸. Ainsi, « je* » *dans un hémisphère* fait un calcul, alors que « je* » *dans l'autre hémisphère* en fait un autre, puis « je » me réunis après dix minutes. Il ne s'agit pas alors d'un « nous », car il y a un engagement à réaliser une unité d'ensemble entre mes « je* » qui n'existe pas au « nous », mais qui implique seulement une division de la conscience en un instant.

C'est dans le sens défendu par C. Rovane qu'il y a un seul moi lorsque je sépare ma conscience en deux pour effectuer deux séries de calculs simultanément. S'il n'y a pas deux êtres, mais un seul, c'est parce que chacune des expériences conscientes séparées s'inscrit dans une unité durable rendue possible par une mémoire et une poursuite d'intention communes. Ces flux de conscience n'ont pas d'intérêts propres et ne cherchent pas à survivre. Ils ne sont que les *instruments* d'une logique à plus long-terme et c'est à cette logique que « je » correspond. Il n'y a qu'un seul « je », car la division a comme finalité la préservation de l'unité d'ensemble de l'être qui persiste dans le temps.

L'exemple du cahier de bord est pertinent. C. Rovane donne un exemple de moyen public d'autoréférence à long terme. Or, la communication entre « je* » *dans un hémisphère* et « je* » *dans l'autre* se fait précisément de cette manière. L'autoréférence dans cette situation n'est donc qu'une extension d'un mode couramment employé dans le temps afin de mener à terme des projets.

Cette remarque montre que définir un moi comme une entité qui s'approprie des expériences sous un mode privé est inadéquat, car c'est avant tout une entité qui dure. Pour durer, elle emploie fréquemment un mode public d'autoréférence. L'imaginaire

²⁴⁷ C. Rovane (1990), p.374-380.

²⁴⁸ C. Rovane (1990), p.368-369, 381-383.

de D. Parfit suppose simplement que ce mode est employé entre diverses parties simultanées plutôt que successives de moi-même pendant dix minutes.

4.3.3- L'attribution à soi comme intuition théorique confuse

Selon l'objection précédente, m'imaginer divisant mon flux de conscience en deux semble possible grâce à une confusion entre « je » et « nous ». J'ai avancé une distinction qui permet de la contredire. « Je » n'est pas nécessairement la chose qui s'attribue l'expérience sous un mode privé, car « je » peut aussi être l'entité qui dure dans le temps en utilisant occasionnellement un mode d'autoréférence entièrement public. Cependant, cette précision ne calme pas forcément l'inconfort que nous ressentons à l'endroit du raisonnement de D. Parfit. Dans ce cas, notre attitude s'explique par autre chose qu'une confusion conceptuelle. Dans ce qui suit, j'expliquerai comment cet inconfort découle d'une intuition théorique que nous confondons avec notre expérience. C'est cela qui est menacé dans cette réflexion.

4.3.3.1- L'expérience directe de soi en un instant

Au premier chapitre, nous avons vu comment l'historique du débat entourant l'identité du moi à travers le temps est constitué d'appel à l'imaginaire visant à nous faire découvrir la véritable nature de notre être. R. Descartes, J. Locke, D. Hume, R. Swinburne, B. William, S. Schoemaker, P. Unger, B. Dainton emploient tous cette méthode. Cela est en continuité avec des usages courants du « je » comme une entité purement mentale qui peut survivre à la destruction du corps. Dans tous ces cas, nous nous définissons comme une entité qui reconnaît sa propre existence de l'intérieur du « flux de conscience » et non de l'extérieur. Suivant cette logique, nombreux sont les critiques de D. Parfit qui pointent un aspect de notre expérience vécue comme marquant une référence à soi impossible à définir autrement. L'argumentation consiste ici encore à regarder ce qui se passe « en nous », présupposant que nous découvrirons ce que nous sommes en nous observant directement être. C'est le cas

même si nous ne le faisons pas en tant qu'âme ou chose purement pensante, mais en tant que dimension, mode, forme, aspect, angle, point, centre, orientation, image, effet, vue, propriété, apparence ou autres catégories ontologiques.

L'unité phénoménale des expériences est un de ces prétendus marqueurs directs de notre présence. Voici quelques autres exemples. G. Strawson défend la présence d'un moi ponctuel dans l'expérience consciente en affirmant que les pensées et expériences sont vécues comme des états d'une chose et non des choses en elles-mêmes. Il la décrit aussi comme l'expérience d'une orientation vers un point et d'une relation directe à soi²⁴⁹. D. Ninan la perçoit dans la centration des expériences²⁵⁰. Q. Cassam et M. Ayer estiment que la proprioception présente ce point central de la perception comme étant le corps perçu de l'intérieur et que cela correspond à l'expérience du moi préalable à sa conceptualisation²⁵¹. D. Zahivi, W. Fasching, A. Gilead et D. Kolak défendent sa présence en invoquant le mode d'accès privé, secret, distinct du mode d'accès aux états mentaux d'autrui, en plus d'énumérer eux aussi l'immédiateté, l'absence d'inférence, la dimension interne du contrôle²⁵². P. Stokes cible dans le sentiment d'avoir à répondre de l'expérience la présence de soi²⁵³. S. Blackburn décrit l'unité du moi comme le point de fuite d'une situation imaginée²⁵⁴.

Ces auteurs défendent tous la thèse que, même en l'absence d'une chose de plus que le corps, le cerveau et une diversité de processus mentaux, l'expérience consciente a intrinsèquement la forme d'une référence à soi²⁵⁵. La démarche de D. Parfit la déforme. Même si la perspective de D. Parfit est possible, elle n'informe pas la nôtre²⁵⁶. Cette insistance marquée d'un aussi grand nombre d'auteurs sur la présence

²⁴⁹ G. Strawson (1999), section 1, 9 et 10.

²⁵⁰ D. Ninan (2009).

²⁵¹ M. Ayers (1991), p.287. Q. Cassam (1997), p.51.

²⁵² D. Zahivi (2000), W. Fasching (2009), A. Gilead (2008), D. Kolak (2008).

²⁵³ P. Stokes (2010).

²⁵⁴ S. Blackburn (1997)

²⁵⁵ Ils diffèrent lorsqu'il s'agit de la préciser.

²⁵⁶ G.Strawson (1999), p.114 lui adresse cette objection dans ces termes.

nécessaire d'un soi dans l'expérience vécue suggère que c'est cela qui résume l'intuition commune. *Nous sommes capable de nous observer directement être et c'est ce qui fait notre spécificité.* Or, c'est ce que D. Parfit défie en nous imaginant ayant deux flux de conscience en même temps.

Comment ? En illustrant une situation où « je » observe directement en un instant *moins* que « moi » dans l'expérience consciente. Lorsque « je » m'observe, « moi », agir de l'autre côté de mon champ de vision, « je » m'observe par l'intermédiaire d'une théorie. Je suppose que les signes perçus renvoient à une activité consciente et qu'elle sera éventuellement intégrée à celle dont « je » ai immédiatement conscience. En ce sens, l'objet d'une perception directe n'est pas « moi », mais simplement une expérience ponctuelle. Son intégration dans un tout plus grand à laquelle se réfère « moi » est médiatisée et non directe. L'imaginaire de D. Parfit l'illustre. Nous pouvons alors comprendre l'unité des expériences en un instant sans l'attribuer à soi, mais à *moins* que soi. Sa supposition, lorsqu'il avance de tels puzzles, c'est que les cas normaux ne sont pas différents, mais seulement plus confus²⁵⁷. Dans les cas normaux aussi, je n'observe pas directement « moi », mais *moins* que « moi ».

Mon usage des guillemets a une fonction, car l'utilisation du « je » et du « moi » peut sembler contradictoire. Si « je » n'observe pas « moi » directement, qu'est-ce qui le fait²⁵⁸ ? Ne suis-je pas en train de reconnaître dans mon expression que c'est « je » qui se reconnaît ainsi ? La réponse de D. Parfit est de substituer « cette » à « je »²⁵⁹.

²⁵⁷ "These assumptions, as I have said, cover actual cases, and our own lives. But they are best revealed when we consider the imaginary problem cases. (...) In ordinary cases, questions about identity have answer. In such cases, there is a fact about personal identity, and Reductionism is one view about the kind of fact this is. On this view, personal identity just consists in physical, and/or psychological continuity. We may find it hard to decide whether we accept this view, since it may be far from clear when one fact consists in another. We may even doubt whether Reductionists and their critics really disagree." D. Parfit (1995), p.27.

²⁵⁸ H. Noonan (1989), chapitre 4, présente cette objection à l'approche de D. Hume. C'est aussi le raisonnement de R. Descartes (1979), même si leurs conclusions diffèrent.

²⁵⁹ "Another ground is something given for belief in such Egos. It can be claimed against any wholly objective description of reality – any description not made from a point of view – that there are certain truth which it omits. One example of such truth is that I am I, or that I am Derek Parfit. (...) Such

Suivant cette logique, « je m'observe directement » renvoie plutôt à « cette auto-observation est directe ». Or, cette auto-observation révèle la présence de *moins* que moi et surtout de *moins* qu'une chose pensante qui dure.

C'est probablement cette conclusion qui frustre nos attentes relativement à l'usage du « je » et non la présence d'une nécessité conceptuelle. Les exemples mentionnés précédemment l'illustrent. Invoquer une orientation du mental (un centre, une occupation d'un point de fuite, etc.), un positionnement corporel, une union phénoménale, une réalité cachée au reste de l'univers et autres traits de l'expérience consciente en un instant est inapte à révéler directement notre présence, car ils révèlent directement *moins* que nous-mêmes. Cette quête d'un soi *immédiat* et *direct* est vaine. Ce raisonnement n'est que l'extension de celui de D. Hume, qui désespère de se trouver lui-même parmi ses perceptions, buttant toujours sur autre chose²⁶⁰.

4.3.3.2- Théorisation du vécu

Après avoir discuté de la possibilité empirique d'une division phénoménale de l'expérience en un seul et même être, nous avons vu une objection conceptuelle. « Je » se définit comme une entité capable de s'approprier des expériences et de se réfléchir sur un mode privé. Or, la fiction avancée par D. Parfit pour montrer comment l'unité des expériences peut être décrite sans se référer à soi ne respecte pas ce sens. Nous avons ensuite vu comment un « je » défini comme unité durable usant parfois d'un mode public contredit cette affirmation. J'ai avancé que le véritable défi posé à nos intuitions n'est pas une confusion conceptuelle, mais la remise en cause de

truth can be stated by Reductionist. The word "subjective" is misleading. What are called subjective truths need not involve any subjects of experiences. A particular thought may be self-referring. It may be the thought that this particular thought, even if exactly similar to the other thoughts that are thought, is still *this* particular thought – or this particular thinking of this thought." D. Parfit (1984), p.252.

²⁶⁰ D. Hume Traité de la nature humaine, Livre I : De l'entendement, 4ème partie, section VI : « Pour ma part, quand j'entre le plus intimement dans ce que j'appelle moi-même, je bute toujours sur quelque perception particulière ou sur une autre, de chaud ou de froid, de lumière ou d'ombre, d'amour ou de haine, de douleur ou de plaisir. Je ne peux jamais, à aucun moment, me saisir moi-même sans une perception, et jamais je ne puis observer autre chose que la perception »

l'intuition selon laquelle « je » reconnaît directement mon être dans l'expérience consciente. Cette explication demeure incomplète. Pourquoi nous attendons-nous à nous observer ainsi ?

L'approche de D. Dennett sur la conscience est éclairante. Selon lui, nous plaçons dans l'expérience consciente plus que ce qu'elle contient en confondant notre vécu avec une théorie sur notre vécu²⁶¹. Il l'illustre avec l'exemple d'un mur supposément rempli de figures de Marilyn Monroe selon un témoignage, alors que le mécanisme causal de la vision indique qu'il n'y a qu'une image de Marilyn située sur le point focal de la vision. Plutôt que de supposer la présence d'une reconstruction visuelle dans la tête d'une image de multiples Marilyn, le plus simple est de conclure que la personne théorise faussement en elle la présence de ces images²⁶². C'est la même chose qui se passe avec l'intuition de vivre un flux ininterrompu de conscience. Nos yeux font des saccades et nous ne percevons alors rien entre elles. Cela nous incite erronément à théoriser sur la présence en nous d'un flux ininterrompu²⁶³.

Il en va peut-être ainsi de toutes les expériences phénoménales de soi. Nous prétendons y voir directement un signe de nous-mêmes à travers un marqueur comme nous prétendons voir des Marilyn partout sur le mur ou un flux ininterrompu de conscience. Nous confondons alors notre vécu et nos théorisations²⁶⁴. Les désaccords quant à ce que cette méthode introspective nous dévoile sont révélateurs.

²⁶¹ D. Dennett (1991), p.346.

²⁶² D. Dennett (1991), p.335: « Your brain just somehow represents *that* there are hundreds of identical Marylins, and no matter how vivid your impression is that you see all the details, the details are in the world, not in your head. And no figment gets used up in rendering the seeming, for the seeming isn't a rendering at all, not even as a bit-map. »

²⁶³ D. Dennett (1991), p.355-356.

²⁶⁴ Je n'affirme pas que la pensée de D. Dennett corresponde à celle de D. Parfit. Ce que j'affirme, c'est que l'expérience consciente *de soi* est illusoire de la manière dont D. Dennett estime que l'expérience consciente est illusoire *en général*.

J. Ganeri s'oppose à D. Parfit d'une manière qui illustre comment cette présence de soi peut être une théorisation. Selon lui, les expériences multimodales combinant différents sens (vision, audition, odorat et autres) témoignent d'une chose intégrative²⁶⁵. Forcément, il y a *une* chose (pas plusieurs) qui intègre ces composants distincts ensemble. Il suppose qu'autrement l'unité devrait être inférée par des raisonnements à propos du monde extérieur et que l'identité entre le sujet des différents aspects de ces expériences pourrait être erronée. Or, cette unité et identité sont immédiatement données et immunisées contre l'erreur²⁶⁶. La référence à soi (et non au cerveau) sert donc d'hypothèse pour expliquer l'unité des expériences et n'est ni l'expression d'une nécessité conceptuelle, ni l'objet d'une introspection.

Cette ligne d'argumentation ressemble à celle qui prouve l'existence de Dieu par la complexité du monde. Le monde est trop éblouissant pour croire qu'une interaction de particules élémentaires stupides puisse le générer, donc il doit y avoir un être supérieur au-delà qui en soit la source. De la même manière, la présence d'un « effet que cela fait d'être » et l'unité de ce phénomène nous semblent trop spéciales pour que des neurones, des axones, des dendrites et divers produits chimiques en interaction puissent les produire. C'est à travers cette théorie que nous postulons que « je » doit être constamment et entièrement présent dans l'immédiat et non étendu en plusieurs morceaux distincts dans le temps et dans l'espace sous la forme d'un agrégat. C'est ainsi qu'H. Langsam suggère que le non réductionnisme visé par D. Parfit est une théorie du sens commun pour expliquer la présence de la conscience²⁶⁷.

²⁶⁵ J. Ganeri (2000).

²⁶⁶ J. Ganeri (2000), p.647-653.

²⁶⁷ H. Langsam (2001), p.255-256: « (...) on Reductionist view, the self is a "terminological posit", not an explanatory posit. By contrast, the Non-Reductionist, the self of folk psychology, is an explanatory posit. (...) Whereas the Reductionist self is posited to indicate that a sufficient number of relations of certain kinds hold between events of certain kinds, the Non-Reductionist self is posited to elucidate a certain intrinsic feature of certain kind of events, their "consciousness". »

M. Siderits objecte à ce genre d'argumentation que l'action de cette chose, pour être comprise par introspection, doit être définie comme une conscience pure détachée de l'expérience elle-même, mais agissant sur elle. Or, dans ce cas-là, son action sur l'expérience apparaît comme étant mystérieuse, c'est-à-dire inexplicable. C'est l'équivalent de faire intervenir un « abracadabra » indéfinissable. C'est une théorie magique sur l'expérience et non une expérience²⁶⁸.

L'imaginaire d'un moi au-delà du corps, du cerveau et de processus mentaux particuliers projette ce construit théorique sur nos pensées comme plusieurs images de Marilyn sur un mur. Nous ne l'observons pas directement, ni en tant que chose pensante, ni en tant qu'apparence, mais en théorisons la présence lorsque nous répondons à des dilemmes philosophiques sur notre vécu. C'est relativement à cette tendance qu'un moi capable d'avoir deux consciences simultanées semble bizarre, car il est impuissant à produire le miracle que nous nous attendons de lui : unir la diversité du réel sous une seule et même apparence. Si nous reconnaissons qu'il n'a jamais ce pouvoir (il le doit à ses composantes), alors nous devrions réviser notre jugement relativement à la bizarrerie de cet imaginaire. La réalité courante ne s'en distingue pas. Notre réticence à le faire en invoquant des motifs de nécessité conceptuelle ou une phénoménologie inévitable de la référence à soi voile le fait que nous avons encore des doutes à cet égard, mais que nous sommes confus sur ce que nous pensons réellement.

4.4. Résumé

En résumé, D. Parfit soutient que notre existence consiste en l'existence d'un corps et d'un ensemble de processus mentaux. Des auteurs s'y opposent en percevant dans l'unité phénoménale un signe que nous sommes une chose supplémentaire au-delà de ces composantes. Afin de contredire cette intuition, D. Parfit en appelle à des cas de

²⁶⁸ M. Siderits (2003), p.18-20.

division de la conscience suite à une césure du corps calleux et à une fiction de division de soi. Dans ces situations, l'unité des expériences doit être décrite en pointant l'état des expériences en elles-mêmes et non l'être qui les vit.

Des objections lui ont été adressées quant à l'interprétation des cas empiriques et quant à l'intelligibilité des concepts employés. Les patients ayant subi une césure du corps calleux ne manifestent pas une division de l'activité consciente assez complexe et durable pour conclure en la présence de deux expériences distinctes. De plus, ils ne démontrent pas les signes d'une division simultanée de leur conscience, mais plutôt d'une brève alternance hémisphérique. D'autre part, il définit dans sa fiction le « je » comme une entité qui renvoie à elle-même essentiellement par des moyens publics, alors qu'il le fait essentiellement via une intériorité. Il n'imagine donc que deux êtres se partageant un corps et non un être possédant deux intériorités.

Or, même si les patients au cerveau divisé ne démontrent pas une division complexe, ils en montrent assez pour faire de la division phénoménale une explication intuitive. De plus, la brève alternance inconsciente de deux expériences ne se distingue pas assez d'une division simultanée pour bloquer les conclusions de D. Parfit. Cette hypothèse suggère la présence d'une chose qui alterne afin de préserver une intuition et non d'enrichir l'explication. L'acharnement à nier la division de la conscience dans ces cas suggère la présence d'une intuition non réductionniste.

D'autre part, « je » peut renvoyer à une unité durable qui adopte occasionnellement un mode public d'autoréférence et se divise. Dans ce cas, le scénario de Parfit est concevable. L'appel à une confusion conceptuelle dissimule un inconfort qui a deux explications concurrentes : la conviction illusoire d'observer directement notre être et une théorisation sur l'apparition de la conscience confondue avec l'expérience consciente elle-même.

CHAPITRE IV

L'IMPORTANCE

Dans les chapitres précédents, nous avons discuté d'aspects de l'expérience consciente associés à la présence d'un moi : l'intuition d'une présence interne, l'autorité sur son passé, l'intégration active dans l'histoire d'une vie et l'unité phénoménale. Selon D. Parfit, nous y voyons un signe que notre existence consiste en une chose supplémentaire au-delà du corps, du cerveau et d'un ensemble de processus mentaux inter-reliés. Afin de faire ressortir la part d'illusion que nous y mettons, il montre comment ces aspects peuvent être compris adéquatement sans se référer à soi et comment les oppositions philosophiques concernant la nature de notre existence sont indéterminées. Relativement à l'existence du corps et d'un ensemble de processus mentaux inter-reliés, notre existence n'ajoute rien de plus qu'une « façon de parler ». Les débats philosophiques sur sa nature véritable se réduisent à une querelle verbale à propos de laquelle il n'y a pas de « fait en la matière ».

Dans ce chapitre-ci, nous discuterons d'une dernière facette de l'expérience consciente qui semble impossible à comprendre sans se référer à l'identité du moi à travers le temps : l'importance²⁶⁹. Qu'un être soit capable de sentir son existence est intuitivement ce qui lui donne de la valeur. C'est pourquoi les roches n'en ont pas. C'est aussi pourquoi les phases de la vie ne possèdent pas d'intérêts propres, car elles ne sont pas conscientes d'elles-mêmes. C'est l'être qui dure en moyenne 80 ans qui l'est. Souffrir maintenant pour jouir demain n'est pas un don ou un sacrifice, mais un acte intéressé compensé lorsqu'un qu'un même être vit les deux moments. Un monde contenant ou non notre existence fait une énorme différence.

Ces attitudes sont difficiles à justifier autrement qu'en s'orientant sur nous-mêmes, puis en contemplant directement comment nous importons. Renvoyer plutôt aux

²⁶⁹ D. Parfit emploie le terme « importance », mais le terme « valeur » pourrait être employé.

composantes qui accompagnent notre existence tend à détruire cette magie. L'énoncé « je survis » n'entraîne pas les mêmes émotions que « cette expérience-ci sera poursuivie via un ensemble de souvenirs, de pensées, de projets, de traits grâce au système complexe de particules physiques qui la supporte en ce moment ».

Si les deux désignent la même réalité, la magie entourant la valeur de mon être est un attachement envers une vue de l'esprit. Des auteurs y voient un argument contre le réductionnisme de D. Parfit²⁷⁰. Il y répond de deux manières. Soit une partie des composantes entourant notre existence possède cette importance et nous pouvons donc la décrire sans renvoyer à soi²⁷¹, soit une partie de l'importance entourant notre existence est sans fondement²⁷². Cependant, il n'est pas rationnel d'accorder l'importance directement à soi.

Dans ce qui suit, je préciserai son argumentation afin de mieux clarifier les enjeux qu'il soulève. Il s'appuie sur une analyse d'une fiction de duplication de soi dont il généralise la conclusion à la vie courante et sur un présupposé selon lequel les éléments constitutifs d'un tout ont plus de valeur que le tout lui-même. Des objections montrent les limites de ce raisonnement : cette fiction est une exception et le tout peut avoir plus de valeurs que ses parties. Ceux qui les formulent préconisent d'évaluer l'importance de notre existence en la situant dans le contexte de la vie courante et des pratiques qui l'entourent et non à partir de sa composition métaphysique. Plusieurs répondent pourtant en faveur de D. Parfit en se servant d'un tel contexte et en élargissant les considérations à prendre en compte dans son raisonnement. J'y affirmerai qu'une approche centrée sur la psychologie et la culture ambiante peut appuyer ses conclusions. Cependant, cette discussion néglige la

²⁷⁰ B. Butler, dans J. Perry (1975), p.102, G. Madell (1981), p.110 et R. Swinburne (1973), p.286.

²⁷¹ D. Parfit (1984), p.255-266.

²⁷² Dans D. Parfit (1984), p.307-312, il présente cette thèse comme « l'affirmation extrême » selon quoi en l'absence d'un fait non-réductible à propos de notre existence, la continuité de notre psychologie n'a pas d'importance. Dans D. Parfit (1995), p.44, il affirme simplement qu'une partie de l'importance associée à cette continuité est illusoire.

présence d'une intuition forte selon laquelle la valeur est l'essence de notre existence d'une manière qui transcende ce genre de justifications. La réflexion de D. Parfit vise à lui reconnaître une part de vérité, mais à lui soustraire une part d'illusion.

5.1- « L'identité n'est pas ce qui importe dans la survivance »

5.1.1- Confusion inhérente au problème de l'identité du moi

Nos intuitions à propos de l'identité du moi à travers le temps portent sur deux considérations pouvant être confondues : ce en quoi notre identité consiste et ce qui lui donne de l'importance. Parmi les attitudes qui expriment la valeur de l'identité du moi à travers le temps, il y a l'anticipation des expériences futures²⁷³, la compensation d'un effort effectué précédemment²⁷⁴, la responsabilité pour une action passée²⁷⁵, le souci de survivre²⁷⁶ et tous les intérêts centrés sur soi plutôt que sur les autres²⁷⁷.

J. Locke fait consister notre existence en une conscience à travers le temps plutôt qu'en une substance parce qu'elle justifie les attitudes qui l'entourent. Il lui apparaît étrange de rendre responsable un être d'un crime dont il n'a pas conscience parce que l'âme ayant causée ce crime aurait transmigré de corps et supporterait sa pensée²⁷⁸. Les approches néo-lockéennes centrées autour de la continuité d'une psychologie raisonnent ainsi. Si je ne suis pas simplement un organisme biologique pensant, c'est parce que la capacité de continuer mon expérience par des souvenirs, des actions, des traits et des croyances est une raison de m'intéresser à moi-même, alors que l'extension de mon corps dans le temps n'en est pas une²⁷⁹.

²⁷³ D. Parfit (1984), p.305-312.

²⁷⁴ D. Parfit (1984), p.329-344.

²⁷⁵ D. Parfit (1984), p.323-326.

²⁷⁶ D. Parfit (1984), p.261-273.

²⁷⁷ D. Parfit (1984), p.282-283.

²⁷⁸ J. Locke, Chapitre 27, section 20.

²⁷⁹ D. Schoemaker (2008), section 2.1, le résume ainsi : « This criterion of identity (and its variations) has been taken to fit particularly well with our practical concerns, both self-regarding and other-regarding. For instance, what seems to matter for self-concern and rational anticipation is that my *psychological* life continue. Anticipation and self-concern are psychological states or attitudes, as are their objects (my future experiences), so a theory that makes those states mine in virtue of their intrinsic relations to one another seems initially quite plausible. »

Cependant, l'objection va dans les deux sens. B. William, T. Nagel ou P. Unger ont l'intuition qu'ils continueraient d'être intéressés au futur de façon personnelle même après une rupture radicale de la conscience d'eux-mêmes à travers le temps du moment que leur corps ou leur cerveau continuerait de supporter une pensée²⁸⁰.

Ces désaccords à propos de ce qui justifie nos attitudes envers nous-mêmes expliquent les désaccords à propos de ce en quoi notre survie consiste. Si une personne croit qu'elle est constituée essentiellement d'une psychologie et non d'un corps, c'est parce qu'elle adopte ces attitudes essentiellement envers les états mentaux futurs qui s'inscrivent en continuité de son expérience actuelle. Si une autre personne croit qu'elle est plutôt constituée d'un corps et non d'une psychologie, c'est parce qu'elle le fait plutôt envers les états mentaux qui impliquent son corps en fonction, sans considération pour leur continuité. Dans ce cas, ce n'est pas ce en quoi consiste notre existence qui justifie nos attitudes, mais ce sont nos attitudes qui justifient ce en quoi consiste notre existence.

D. Parfit intervient dans ce débat en opposant deux thèses.

Selon la première, c'est l'identité du moi à travers le temps qui justifie l'importance morale et rationnelle des faits qui l'accompagnent. Par exemple, confronté au scénario d'une destruction de notre corps suivie d'une réplication parfaite, c'est notre croyance selon laquelle la personne répliquée serait un autre être qui nous amène à ne pas adopter à son endroit une attitude comparable à celle que nous avons envers nous-mêmes²⁸¹. L'importance de notre continuité psychologique dérive ainsi de l'identité

²⁸⁰ P. Unger (1990), B. Williams (1970), T. Nagel (1986).

²⁸¹ D. Parfit (1984), p.279-280: « On the Reductionist View, my continued existence just involves physical and psychological continuity. On the Non-Reductionist View, it involves a further fact. (...) When I fear that, in Teletransportation, I shall not get to Mars, my fear is that the abnormal cause may fail to produce this further fact. »

du moi à travers le temps et non l'inverse. Il résume cette thèse dans l'expression « l'identité est ce qui importe dans la survivance »²⁸².

Selon la seconde thèse, ce sont des faits accompagnant l'identité du moi à travers le temps qui lui donnent son importance et non l'identité elle-même. Dans la réplique, nous pouvons évaluer que l'interruption de la continuité physique, du potentiel de conscience et l'anormalité des causes continuant notre psychologie nous enlèvent les raisons de traiter l'être résultant comme nous-même²⁸³. Dans ce cas, « l'identité n'est pas ce qui importe dans la survivance. » L'identité dérive alors de l'importance que nous accordons à autre chose qu'elle²⁸⁴.

Selon D. Parfit, notre réaction aux problèmes de l'identité du moi à travers le temps indique une intuition favorable à la thèse « l'identité est ce qui importe »²⁸⁵. Selon lui, l'absence d'un fait de plus à la continuité du corps et d'une psychologie à travers le temps supporte la thèse « l'identité n'est pas ce qui importe »²⁸⁶.

5.1.2- Le paradoxe de la duplication

Il développe son argumentation en analysant les paradoxes posés par la fission.

²⁸² D. Parfit (1984), p.215 : « In one of the outcomes, I am about to die. In the other outcome I shall live for many years. If these years are worth living, the second outcome would be better for me. And the difference between these outcomes would be judged to be important on most theories about rationality, and most moral theories. (...) What is judged to be important here is whether, during these years, there will be someone living who will be me. (...) On one view, this is always what is important. I call this view that personal identity is what matters. »

²⁸³ D. Parfit (1984), aux p.282-287 discute ainsi des approches réductionnistes centrées sur une chose physique comme devant fonder les attitudes entourant notre existence sur les composantes de cette chose en elle-même sans faire référence à la présence de soi.

²⁸⁴ D. Parfit (1984), p.215.

²⁸⁵ Il traite les fictions problématisant l'identité comme une méthode pour découvrir ce que nous croyons. D. Parfit (1984), p.200. Il affirme ensuite de la thèse de l'importance de l'identité : « This is the natural view ». D. Parfit (1984), p.215.

²⁸⁶ D. Parfit (1984), p.261-266.

À l'origine, B. Williams invoquait à l'endroit de J. Locke le problème de la duplication. Si c'est la conscience de soi qui constitue notre identité, alors deux êtres pourraient la poursuivre et être nous, ce qui est contradictoire. C'est le cas si une réplique de nous-mêmes était créée²⁸⁷. Si nous introduisons une clause de non-embranchement stipulant que la conscience de soi doit être unique²⁸⁸, alors la création ailleurs d'un être ayant une conscience de notre passé peut nous tuer. Cette conséquence est absurde, car ce fait est extrinsèque à la relation entre les parties de nous-mêmes dans le temps. Qu'une réplique soit créée ou non, la réalité désignée en pointant du doigt le corps original restera la même. La duplication de sa conscience lui est étrangère. Y voir la fin de notre existence la fait donc dépendre d'un fait apparemment trivial. Cette conclusion est invraisemblable²⁸⁹.

R. Swinburne a retourné cet argument contre toutes les approches qui font dépendre l'identité du moi à travers le temps d'autres faits plus particuliers, que ce soit la continuation d'une chose physique ou d'une psychologie. Au même titre que la psychologie d'un moi peut être logiquement dupliquée, son corps peut l'être aussi. C'est le cas lors d'une double transplantation de moitié de cerveau munie de capacité égale de continuer notre psychologie, scénario recevant le nom de fission. Dans ce cas-là, mon existence dépend aussi d'un fait extrinsèque à la relation entre mes parties. La simple transplantation d'une moitié de cerveau suffirait à me sauver. Donc le succès ou l'échec de l'autre transplantation devient un fait extrinsèque comparable à la création d'un être muni de ma conscience. R. Swinburne en conclut que le moi ne peut être qu'une chose simple et indivisible²⁹⁰.

²⁸⁷ B. Williams (1973), « personal identity and individuation ».

²⁸⁸ C'est la solution de S. Shoemaker (1984).

²⁸⁹ B. Williams (1973), p.20.

²⁹⁰ R. Swinburne (1984), p.13-18.

5.1.3-La solution de D. Parfit

D. Parfit en tire une autre conclusion. Bien que plusieurs descriptions de la fission soient possibles, celle qui pose qu'une personne est remplacée par deux nouvelles lui apparaît être la meilleure. « The best description is that neither of the resulting person would be me²⁹¹. » Il n'y aurait aucune raison d'être une personne plutôt que l'autre, car les deux auraient des raisons égales de prétendre être nous²⁹². Nous pourrions à exister à deux places en même temps en suivant le modèle de l'examen de physique (voir le 4^{ème} chapitre), mais cela modifie considérablement le sens d'une personne²⁹³. Cependant, si le scénario est problématique, c'est parce que nous supposons que notre existence dans le futur est ce qui importe. Il est étrange que ce qui nous importe tant dépende d'une relation extrinsèque. Pour résoudre cette dissonance, il suffit de substituer la thèse selon laquelle notre existence importe par celle selon laquelle c'est plutôt la continuation de notre psychologie et/ou d'une partie suffisante de notre corps. Cette relation dépend alors d'un fait intrinsèque²⁹⁴.

Le raisonnement de D. Parfit est que l'existence du moi par rapport à la continuation d'une psychologie et/ou d'un corps n'est que l'addition d'une « façon de parler ». Or, les faits réels sont plus importants que le choix de mots pour les présenter. En ce sens, les éléments de notre existence doivent contenir l'essentiel de notre valeur et non notre existence en elle-même parce que celle-ci ne leur ajoute rien de substantiel. Cela est vrai dans les cas normaux, mais les situations imaginaires interviennent comme un moyen d'en illustrer le principe. Dans ce cas, la *réalité* de notre relation

²⁹¹ D. Parfit (1984), p.260.

²⁹² D. Parfit (1984), p.256.

²⁹³ D. Parfit (1984), p.256 : « There remains a fourth possibility : that I survive as both of the resulting people. This possibility might be described in several way. I might first claim "what we have called "the two resulting people" are not two people. They are one person. I do survive this operation. Its effect is to give me two bodies, and a divided mind". This claim cannot be dismissed outright. As I argued, we ought to admit as possible that a person could have a divided mind. (...) But though this description of the case cannot be rejected as inconceivable, it involves a great distortion in our concept of a person."

²⁹⁴ D. Parfit (1984), p.261-273.

au futur est inchangée et seul le *nom* pour en parler change. Or, cette différence est tellement mince que lui accorder beaucoup d'importance apparaît irrationnel²⁹⁵.

5.1.4-Implications

La première conséquence de ce changement de perspective, selon lui, c'est que la continuation de notre corps perd de son importance. Si l'identité du moi est mise de côté, la transplantation d'un cerveau dans un corps similaire n'est qu'un moyen de poursuivre son activité mentale au moyen de multiples transplantations d'organes en même temps. Décrit ainsi, cela ne peut être tragique²⁹⁶. Intuitivement, c'est le cerveau et non le corps qui nous importe, car il est la cause de notre activité mentale. Or, l'enregistrement des informations sur l'état du corps, la destruction du corps, l'envoi des informations sur Mars et la création d'une copie exactement similaire là-bas ne seraient qu'une façon *anormale* de produire le même résultat²⁹⁷. Selon D. Parfit, c'est l'équivalent pour un aveugle de récupérer la vue au moyen d'un appareil artificiel ; sauf qu'il s'agit ici de recouvrer la vie²⁹⁸. Cela n'a rien de dramatique, car c'est l'effet qui nous importe et non la manière dont il est amené²⁹⁹.

La deuxième conséquence de ce changement de perspective, c'est que les attitudes centrées autour de soi peuvent changer de portée et perdre en intensité.

D'abord, la continuation de notre psychologie a deux composantes : la connectivité et la continuité. Par « connectivité », il entend un lien direct entre deux événements mentaux d'une vie. Par exemple, la mémoire est un lien direct à une expérience passée. Par « continuité », il entend la présence d'une chaîne reliant ces connectivités

²⁹⁵ D. Parfit (1995), p. 33: "We are discussing cases where, relative to the facts of some lower level, the higher-level fact is, in the sense that I have sketched, merely conceptual. My claim is that such conceptual facts cannot be rationally or morally important. What matters is reality, not how it is described."

²⁹⁶ D. Parfit (1995), p.34-35.

²⁹⁷ D. Parfit (1984), p.284.

²⁹⁸ D. Parfit (1984), p.285.

²⁹⁹ D. Parfit (1984), p.286.

à d'autres. Ainsi, je peux me souvenir d'une expérience ayant eu lieu dans ma vingtaine et ne pas me souvenir d'une expérience ayant eu lieu dans mon enfance, alors que, dans la vingtaine, je me souviens de cette expérience dans mon enfance. Cette chaîne me relie présentement à ce souvenir d'enfance³⁰⁰. Selon Parfit, si l'identité du moi à travers le temps n'est pas ce qui importe, alors il est raisonnable de proportionner les attitudes associées habituellement à soi (anticipation, compensation, attribution de responsabilité, sentiment de survivre) à la présence des connectivités au mépris de la continuité³⁰¹. La raison intuitive de ne pas le faire, c'est que le futur au-delà de ces connectivités est le nôtre. Cependant, si l'identité n'importe pas, ce raisonnement n'a plus de poids³⁰². Autrement dit, qu'un état mental dans 60 ans s'inscrive dans un ensemble d'expériences et de pensées largement étrangers à celui qui constitue notre instant présent, mais en découle peut raisonnablement nous indifférer. Or, que ce futur soit le nôtre n'ajoute qu'un mot différent pour parler de la même réalité, donc cela ne constitue pas une objection substantielle.

Ensuite, ces attitudes perdent en intensité. Le fait qu'une expérience future soit psychologiquement connectée à ceci maintenant apparaît moins important que le fait qu'elle soit la nôtre. Or, si le fait qu'elle soit la nôtre n'ajoute qu'un mot de plus pour la présenter, le surplus d'émotions avec lequel nous lui donnons de l'importance est

³⁰⁰ D. Parfit (1984), p.204-209, p.298-306.

³⁰¹ D. Parfit (1984), p.313: « My concern for my future may correspond to the degree of connectedness between me now and myself in the future. Connectedness is one of the relation that give me reasons to be specially concern about my own future. It can be rational to care less, when one grounds for caring will hold to lesser degree. »

³⁰² D. Parfit (1984), p.315: « A Self-Interest theory might appeal to *The truism*: All parts of a person future are equally parts of his future. (...) This argument assume that personal identity is what matters. (...) Nothing is shown by an argument with a false premise. (...) This would be a good argument if the Non-Reductionist was true. On that view, the truism is a profound truth, deep enough to support the argument. But on the Reductionist view, the truism is too trivial to support the argument. »

irrationnel³⁰³. Une différence réelle aussi mineure ne peut justifier une différence d'attitude aussi majeure³⁰⁴.

5.1.5- Résumé

Les problèmes entourant l'identité du moi à travers le temps servent à justifier un ensemble d'attitudes incluant le désir de survivre, la compensation des efforts, la responsabilité et l'anticipation du futur. D. Parfit oppose à cet effet deux thèses, l'une qui accorde de l'importance à l'identité elle-même et l'autre qui l'accorde à une partie de ses constituants. Son argumentation tire son origine de l'objection de duplication adressée par B. Williams à J. Locke, puis reprise par R. Swinburne contre toutes les théories qui font consister le moi en des faits plus particuliers. Selon D. Parfit, la duplication montre plutôt qu'en l'absence d'un fait supplémentaire à la continuation d'une chose physique et de la psychologie, il n'y a aucune théorie qui peut défendre l'importance de l'identité elle-même sans la faire reposer sur une relation extrinsèque et triviale. Il en conclut que ce n'est pas l'identité qui importe, mais la relation intrinsèque qui unit deux instants. Cela l'amène à dévaloriser l'importance de la continuité physique, à justifier nos attitudes relatives à soi proportionnellement au degré de connectivité nous unissant au futur et à diminuer l'intensité avec laquelle nous le faisons.

5.2- Les enjeux soulevés

Plusieurs objections ont été soulevées et elles peuvent nous aider à mieux comprendre les éléments de confusion qui restent à préciser dans ce raisonnement. Premièrement, il ne va pas de soi qu'un construit linguistique ou un tout composé de parties ne doit pas importer plus ou avoir préséance sur la valeur des réalités plus particulières qu'il

³⁰³ D. Parfit (1984), p.281-282 et (1995), p.44-45.

³⁰⁴ D. Parfit (1984) emploie le terme « irrationnel » par analogie avec le vertige, p.279. Il (1995), p.28 le fait aussi par comparaison avec un comportement superstitieux voulant qu'une souffrance n'ait pas lieu le mardi plutôt qu'une autre journée. D. Parfit (1984), p.308.

sert à désigner. Par exemple, face à l'éventualité d'une fission, nous sommes contraints d'adopter une perspective à la troisième personne en spécifiant qui est qui et cette altération est ce qui rend la situation dramatique. Deuxièmement, D. Parfit tire une conclusion générale d'un cas particulier, alors que nous pouvons le traiter comme une exception à une règle. Si la fission en devenait une, nous serions incapable de dire ce qu'il advient de nos valeurs. Cette conclusion est renforcée par la présence de réactions différentes dans d'autres contextes.

En réponse à ces objections, la réduction de la valeur d'une entité à celle de ses constituants a un certain poids qui peut devenir déterminant s'il est accompagné d'un ensemble d'autres justifications. Les motivations à se libérer des sentiments négatifs associés à la séparation des personnes peuvent le faire. De plus, des précisions additionnelles concernant les éléments qui nous constituent comme l'habitude à orienter nos projets dans une seule direction ou le fait que nous situons rarement notre être dans notre perspective supportent également les conclusions de D. Parfit. D'autre part, elles ne sont pas en contradiction avec nos comportements observables et nos valeurs ambiantes, car l'ensemble de notre vie constitue rarement le cadre de référence de nos émotions et de nos actes, puis il est utile de prendre plutôt les connectivités psychologiques comme centre de ces attitudes.

5.2.1-Le tout importe-t-il plus que les parties ?

5.2.1.1-Le raisonnement par la base

Le raisonnement de D. Parfit part d'une observation : il existe en nous une différence d'attitudes entre deux niveaux de généralité. Le niveau de description le plus général a une importance considérable, alors que le niveau plus particulier nous inspire davantage d'indifférence. Dans la fission, le premier niveau implique que nous cessons d'exister. Cela nous inspire des sentiments tragiques. Or, le second niveau précise plutôt que l'expérience qui nous constitue en un instant sera intégrée dans

deux ensembles plutôt qu'un. Cela nous apparaît banal, en partie parce que c'est un fait externe à ce qui nous constitue actuellement. Or, il présume que le second niveau d'analyse a logiquement préséance sur le premier.

Cette argumentation n'est pas propre aux problèmes de l'identité du moi à travers le temps. Par exemple, les anarchistes libéraux décrivent la taxation comme une forme de vol. Selon eux, nous jugerions ainsi l'acte de prendre le fruit du travail de quelqu'un sous la menace d'une séquestration. Or, le même geste venant de fonctionnaires de l'État apparaît légitime. La dissonance entre les deux niveaux d'analyse, ainsi que la supériorité du plus particulier sur le plus général, constitue la forme de cette argumentation³⁰⁵.

5.2.1.2-Le raisonnement par le niveau supérieur

Une objection contre ce raisonnement montre que la déduction se fait en sens inverse. Des événements dans un cerveau peuvent n'avoir aucune importance en eux-mêmes, mais en prendre à partir du moment où ils constituent des événements mentaux. C'est alors l'importance du niveau supérieur, l'esprit, qui détermine celle de ses parties³⁰⁶. De la même manière, si la continuité d'un objet physique et/ou la poursuite de processus mentaux reliés singulièrement permet de me constituer, alors c'est leur contribution à mon existence qui leur donne une importance. Plutôt que de résoudre la dissonance entre nos attitudes relatives aux deux niveaux en priorisant le niveau inférieur, nous devons prioriser le niveau supérieur. En ce sens, l'absence d'embranchement de la relation de continuité psychologique ou l'anormalité de ses

³⁰⁵ M. Rothbard (2006) développe ainsi son argumentation et elle est reprise en détail par M. Huemer (2013). Les divers blogues, pages et sites libertariens font appel abondamment à ce raisonnement.

³⁰⁶ M. Johnston (1997), p.167-168 : "Particular ontological reductions for every object, event, state and process in the manifest world will locate for each its own particular patterns of constituting facts in the fundamental realm. Now take any valued object, event, state or process and the fact that it exists, obtains, occurs. That fact will be constituted by facts about micro-physical properties, facts about which one will have no particular non-derivative concern (...) But this is not a proof of nihilism. It is a reductio a absurdum of the argument from below." D. Brink (1997), p.117-118 et B. Garrett (1998), p.90 lui adressent une objection similaire.

causes n'importent pas en elles-mêmes, mais le font à partir du moment où leur contribution à mon existence est spécifiée³⁰⁷.

Que l'addition « j'existe » soit l'addition d'une « façon de parler » sur des faits plus particuliers peut être de la plus haute importance. Les exemples ne manquent pas. « Je t'aime » suggère des émotions différentes de « cet ensemble d'expériences-ci contient un sentiment d'amour envers cet ensemble-là. » Nous n'en concluons pas que l'amour doit importer aussi peu que vis-à-vis cette description³⁰⁸.

Cette conclusion est renforcée par le fait qu'en dehors de la fission, d'autres contextes nous amènent à valoriser l'identité elle-même. Nos réactions intuitives face à la réplique vont dans le sens contraire de la position de D. Parfit. Même s'il a raison d'évaluer comme étant sans importance l'anormalité des causes continuant notre psychologie, notre réticence à traiter la réplique avec la même attitude qu'envers nous-mêmes indique l'importance que l'identité a à nos yeux. Plusieurs n'acceptent pas de considérer ce scénario comme un cas d'indétermination. La perte d'identité leur y apparaît clairement, alors qu'elle demeure indéterminée dans la fission. Cette différence est suffisante pour ne pas évaluer les deux cas de la même manière³⁰⁹. Si la réplique est aussi bonne, alors continuer notre expérience à travers une personne qui lit minutieusement sur notre vie, puis la poursuit le serait aussi ; ce qui apparaît contre-intuitif. Or, tant que nous sommes dans l'évaluation des causes, il n'y a pas moyen de bloquer cette inférence. C'est en se centrant sur l'identité que nos intuitions sont respectées³¹⁰.

³⁰⁷ B. Garrett (1998), p.88-91, M. Johnston (1997), p.154-156, E. Sosa (1990), p.321, D. Brink (1997), p.117-118.

³⁰⁸ S. Wolff (1986) élabore abondamment comment le raisonnement de D. Parfit est ainsi en contradiction avec nos comportements affectifs envers nos proches. M. Johnston (1997), p.164-165 insiste sur ce fait.

³⁰⁹ M. Johnston (1997), p.163-168. P. Unger (1990) affirme la même chose.

³¹⁰ C'est la critique que lui fait H. Noonan (1989), chapitre 9.

De plus, des auteurs nient que la fission préserverait ce qui nous importe, car elle implique un changement radical de perspective. Selon S. Blackburn, la perte de singularité de notre flux de conscience nous empêche de nous imaginer l'effet que cela nous fera de vivre une expérience future. Elle nous contraint de nous la représenter sous un mode à la troisième personne³¹¹. P. Unger y voit une perte de focus, qu'il illustre par notre réticence à se multiplier par mille³¹². Si la qualité des éléments constitutifs d'une vie varie en vertu de leur contribution à l'expression d'UNE personnalité, alors ne plus pouvoir le faire détruit ce qui importe³¹³. Le trait commun entre ces réponses est d'accorder l'importance au point de vue d'ensemble constitué par les processus mentaux et non aux processus eux-mêmes.

5.2.1.3- Raisonnement ayant du poids supporté par d'autres considérations

Une réponse nuancée va dans le sens de D. Parfit. Nos attitudes relativement aux entités situées au niveau inférieur de généralité ne peuvent pas automatiquement être utilisées pour contredire celles que nous adoptons envers les entités du niveau supérieur. Cependant, cet argument a un poids accompagné de justifications supplémentaires.

R. Martin ne croit pas qu'une théorie de l'identité du moi à travers le temps puisse démontrer ce qui *doit* importer dans nos attitudes, mais elle peut faciliter un

³¹¹ S. Blackburn (1997), p.196-198: "I am not here presented empirically, as part of the narration. If we say that the self is fantasized as present in this kind of cases, that is simply through the fact that a first-person perspective is adopted. Treating it as part of the scene because of this would be exactly the same mistake as Parfit's treating first-person aims as merely part of the field of consideration present in deliberation. (...) If practical reasoning is essentially conducted from within first-person perspective, imagining my doing something is a very different matter from imagining someone else doing something. The imagining is quite different, and if what I have said is right, this difference is absolute, and will not succumb to indeterminacies in constitutive questions of identity through time." Il appuie son analyse sur celle de D. Velleman (1996).

³¹² P. Unger (1990), chapitre 8.

³¹³ M. Schechtman (1990), p.87-91.

changement en nous amenant à dévaluer les sentiments négatifs associés à la perception d'être séparé du vécu des autres³¹⁴.

Il est d'accord que l'addition d'un langage personnel sur un ensemble impersonnel importe selon les tempéraments³¹⁵. Cependant, il reconnaît en nous une tendance à dévaluer une entité lorsque nous réalisons qu'elle se distingue d'une autre conceptuellement plutôt que substantiellement. Découvrir qu'il en va ainsi de notre existence peut donc avoir cet impact, surtout s'il y a d'autres motivations à aller dans cette direction³¹⁶.

Dans le cas discuté ici, la motivation est de se libérer des sentiments négatifs associés à notre existence. Il y a une satisfaction à imaginer le futur de notre point de vue, mais il y a aussi une frustration à ne pas ressentir celui des autres. Si la radicalité de la frontière qui justifie ce blocage nous apparaît comme étant une vue de l'esprit, le résultat peut être de jouir davantage de ce qui arrive aux autres³¹⁷. L'intimité avec autrui peut être ressentie par la dissolution de ce qui nous sépare, l'intégration commune dans un ensemble de « connectivités mentales » (souvenirs, projets, traits, croyances) et non la reconnaissance de nos individualités réciproques. La fission et toutes les fictions d'indétermination nous montrent comment réduire ainsi la différence entre « moi » et « un autre » à presque rien.

³¹⁴ R. Martin (1998), p.27-30.

³¹⁵ R. Martin (1998), p.16-24.

³¹⁶ R. Martin (1998), p.24 et 27 : « In fairness to Parfit, I think that many people, if they would reflect carefully on his many clarifying examples and arguments, would decide that physical continuity does not matter as much as they had thought it matters. The reason, I believe, is that for these people such reflection would tend to be transforming and liberalizing; or perhaps, it would help them realize that their values, all along, were more liberal than they thought they were. (...) For most of us, in certain kind of cases, seeing mere linguistic conventions for what they are tends to erode their influence on our values and attitudes. For some of us it tends to erode their influence deeply. And personal identity is one of these kinds of cases ».

³¹⁷ D. Parfit (1984), p.281 : « My life seemed like a glass tunnel, through which I was moving faster every year, and at the end of which there was darkness. When I changed my view, the walls of my glass tunnel disappeared. I now live in the open air. There is still a difference between my life and the lives of other people. But this difference is less. »

5.2.1.4- Précisions charitables sur la base avant de juger

Cette discussion est obscurcie également par une tendance à conclure trop rapidement. Si les éléments constituant notre être (corps, cerveau, mémorisation des expériences, poursuite d'intention, persistance de traits de caractère, etc.) nous inspirent une distance émotive par rapport à notre existence, il est possible que ce soit parce que nous sommes restés vagues sur leur nature.

C. Rovane explique ainsi les réticences à accepter la fission comme un substitut à la survie normale par l'habitude de projeter nos vies de façon unidirectionnelle. Nous agissons en fonction d'un futur. Confrontés à la possibilité d'une duplication, nous sommes face à une situation qui en menace la réalisation. En effet, la présence de deux personnes poursuivant nos projets crée des problèmes de coordination. Si elles grandissaient plutôt en sachant qu'elles seront divisées à un moment donné, elles formeraient des projets en conséquence. Or, ce serait alors la perspective de ne pas se diviser qui causerait cet inconfort. Ce n'est donc pas l'identité elle-même qui explique notre réaction, mais l'incapacité de planifier le futur qui en résulte³¹⁸.

M. Belzer nie quant à lui l'omniprésence d'un acte d'orientation à soi dans la vie. La perspective à la première personne implique de ne pas chercher l'identité de celui qui est dans la scène. S. Blackburn explique ainsi la sensation de perdre ce qui importe dans la fission par la nécessité de spécifier qui fera quoi après l'opération, perdant tragiquement la spontanéité de notre perspective. Or, selon M. Belzer, spécifier qui fera quoi équivaut à spécifier où et quand pour discerner plusieurs projections futures³¹⁹. Nous ne pouvons pas nous imaginer en train de faire deux choses en même temps, mais nous pouvons nous imaginer les deux actions l'une après l'autre. Il nous suffit d'indiquer où et quand elles se feront pour ne pas les confondre. Suivant cette logique, nous ne devons pas spécifier l'identité des êtres après la

³¹⁸ C. Rovane (1990).

³¹⁹ M. Siderit (2003), p.66-71 offre une réponse similaire à S. Blackburn.

fission, mais pouvons simplement nous imaginer successivement les deux séries d'actions entreprises en détaillant au besoin le lieu et le moment. La fission aura lieu et ensuite les deux séries d'actions seront mises en branle³²⁰.

Les remarques de M. Belzer défont une confusion entre la perspective à la troisième personne et sans personne. Face à l'éventualité de doubler notre flux de conscience, nous pouvons y voir deux autres personnes agir par procuration sur nos intentions. Dans ce cas, il s'agit d'une perspective à la troisième personne. Cependant, nous pouvons dès le départ ne voir *personne* dans la situation, ni nous-mêmes, ni un autre. Il n'y a que des intentions futures qui s'enchaînent dans un processus mental, puis qui s'intègrent finalement dans deux ensembles sans qu'à aucun moment une pensée ne soit venue lui accoler une étiquette identitaire. Adopter la perspective sans personne anéantit la « réaction d'unité ». Face à la fission, nous projetons différentes intentions dans le futur et elles vont dans deux directions. À aucun moment nous souffrons de l'absence d'un être, car à aucun moment nous jouissons de sa présence, étant absorbé dans le flux de conscience au point de ne pas s'en distancer³²¹.

C'est ce que nous faisons la plupart du temps. En effet, les philosophes doivent apporter des précisions pour problématiser la fission. Sans elles, les gens n'y verraient probablement aucun problème³²². C'est parce que nous ne nous situons pas dans le futur pour vivre, étant entièrement absorbés dans le flux de conscience.

³²⁰ M. Belzer (2005), p.155-159.

³²¹ M. Belzer (2005), p.162 : « We should distinguish, then, between (a) person-level descriptions of the first-person attitudes involved in agency, and (b) the first-person perspective that is itself involved in agency. Instead of supporting the grip of the unity reaction, Velleman's fine-grain analysis can help us see why the unity reaction does not have an absolute grip on our self-conceptions as agents. In practical reasoning from within the first-person perspective, no personal assumption at all about identity or unity need to be made. But of course these assumptions naturally are made when we describe first-person practical and agency - as in Blackburn's statement that imagining my doing something is different from imagining someone else doing it. »

³²² Plusieurs fictions télévisuelles font intervenir de tels embranchements et les auditeurs vont rarement réagir comme les philosophes. M. Belzer (2005), p.155 donne l'exemple de « retour vers le futur ».

En résumé, des précisions concernant la nature du point à l'œuvre dans la vie montrent comment les processus mentaux eux-mêmes sont ce qui nous importe. L'identité elle-même n'en est qu'un ajout superflu qui nourrit de faux problèmes.

5.2.2-Exception à une règle ?

5.2.2.1-Généralisation à partir d'un cas imaginaire

Nous venons de voir comment le raisonnement de D. Parfit montrant que l'identité n'est pas ce qui importe dans la survivance présume que nos attitudes relativement au niveau inférieur d'analyse doivent avoir préséance sur ceux du niveau supérieur. Cependant, la déduction peut se faire dans l'autre sens. Malgré tout, accompagnés d'autres considérations, nous avons, en général, un biais en faveur du niveau inférieur d'analyse qui peut faire une différence. De plus, l'évaluation favorable envers le niveau supérieur peut résulter d'une compréhension vague des éléments impliqués.

La forme du raisonnement de D. Parfit pose un second problème³²³. Il généralise à la vie courante une conclusion qu'il tire d'une seule fiction. Selon lui, l'identité du moi à travers le temps accompagne ce qui importe dans la survivance et les deux sont confondus. Dans les cas normaux, nous ne décidons jamais si les attitudes relatives à soi se justifient par l'existence du corps, des processus mentaux particuliers ou la présence d'un seul et même être à travers le temps. Afin de s'en faire une idée claire et distincte, D. Parfit imagine une situation où les deux justifications entrent en conflit. Il raisonne comme un scientifique qui isole un ingrédient chimique X d'un composé XY pour voir s'il va produire la même réaction et donc déterminer si le rôle de Y est superflu. Constatant en imagination un cas où l'identité n'est pas ce qui

³²³ L'analyse qui suit s'appuie sur celle de T. Z. Gendler (2002), qui rejoint et résume de façon concise et systématique des critiques adressées par M. Johnston (1997), P. Unger (1990) et S. Wolf (1986).

importe, il en conclut que l'importance accordée à l'identité du moi à travers le temps est toujours superflue par rapport aux éléments qui l'accompagnent³²⁴.

5.2.2.2-Raisonnement étranger au contexte de la rationalité courante

Or, il est plus raisonnable de voir dans la fission une exception à une règle. Dans la vie courante, notre existence future supporte les attitudes relatives à soi. D. Parfit a simplement imaginé une situation où ce n'est pas le cas.

Sa conclusion ne vaut que dans les cas où il y a indétermination de l'identité du moi et non dans ceux où le résultat est clairement une autre personne. Dans la fission, nous pourrions considérer les deux personnes résultantes comme nous-mêmes étant à deux endroits en même temps. Lors de la destruction de notre corps suivi de la création d'une copie parfaite, il n'y a pas d'indétermination : c'est une autre personne. Cette différence explique pourquoi, dans le second cas, nous ne préservons pas nos attitudes, mais que nous le faisons dans le premier. Au lieu de conclure que l'identité du moi à travers le temps n'est pas ce qui importe, il est plus juste d'affirmer qu'elle est ce qui importe et que c'est pourquoi nous étendons cette importance à d'autres êtres dans les cas étranges où sa présence est indéterminée³²⁵.

Cette réponse s'appuie sur le rôle de notre identité dans un ensemble de pratiques courantes plutôt que sur des justifications d'ordre métaphysique³²⁶. Accorder de l'importance à l'identité plutôt qu'à des processus mentaux produit un sentiment d'intimité individuelle avec eux difficile à entretenir envers une série d'événements. Cela permet aussi de veiller au développement des enfants ou d'entreprendre des projets à long-terme en négligeant l'importance des changements qu'ils impliquent³²⁷.

³²⁴ T. Z. Gendler (2002), p.37-44.

³²⁵ M. Johnston (1997), p.166, 167,170. P. Unger (1990), chapitre 7, T. Z. Gendler (2002), p.47-49.

³²⁶ M. Johnston (1997), qualifie cette position de « minimalisme ».

³²⁷ S. Wolf (1986).

Ces impacts justifient le rôle de l'identité. La réponse de D. Parfit n'a de signification qu'en tant qu'exception qui « parasite » cette norme³²⁸.

5.2.2.3- Raisonement en accord avec le contexte courant de justifications

Or, même si le raisonnement de D. Parfit ne garantit pas sa conclusion, nous pouvons montrer comment la vie courante et les justifications qui l'entourent contiennent suffisamment d'éléments qui parlent en sa faveur.

D. Shoemaker répond ainsi que la perte ou la rétention de connectivités comme la mémorisation et la présence ou absence d'un projet unificateur affectent les patterns de distanciation ou d'identification indépendamment de l'identité du moi à travers le temps. L'amnésie est un exemple. La plus grande capacité à s'imaginer de quoi aura l'air le futur joue effectivement un rôle dans nos attitudes d'anticipation³²⁹. Il est faux que le cadre de narration est pour tous la vie en général. Il est souvent moindre. La position de D. Parfit accorde donc la rationalité avec la pratique commune plutôt que de s'y opposer³³⁰.

D'autre part, les conclusions de D. Parfit s'appuient sur des justifications conséquentialistes. M. Siderits considère ainsi que d'accorder de l'importance à l'identité elle-même la fragilise au même titre que faire reposer l'amour sur sa magie. Elle est alors dépourvue de fondement. Pourtant, le système psychologique d'expériences fortement intégrées qui correspond à moi représente l'unité la mieux à

³²⁸ T. Z. Gendler (2002), p.50-51: « (...) suppose we have no sense that there could be personal identity, as opposed to mere R-Relatedness [connectivité et continuité psychologique entre les expériences et les pensées] – would there still be such thing as prudential concern? I suggest the answer is “No”. The concept of prudential concern is tied up with concepts of fairness, responsibility, justice and rationality. Our views about the sorts of rational and moral obligations we have to ourselves and others, considered as beings who exist through time, rest on assumption that each of us will have at most one continuer, and that this continuer is someone with whom we will be identical. Disruption of background assumption would result in disruption of the entire framework by which we make sense of this wide range of concepts. (...) The interpretation of the exceptional case is parasitic on the interpretation of the normal cases.”

³²⁹ D. Shoemaker (1999), 405-409.

³³⁰ D. Shoemaker (2008), section 2.4

même d'influer sur la qualité des expériences. Cela justifie de la prendre comme objet de souci en nous amenant à comprendre pourquoi. Ce système possède des capacités d'autocontrôle, d'auto-révision et d'auto-examen qui en font le meilleur « maximisateur » d'expériences de qualités. Dans ce cas, je me soucie d'un être futur qui continuera ma psychologie parce que c'est l'être à propos duquel j'ai un pouvoir d'influence. C'est le cas qu'il soit moi ou un double de moi résultant d'une fission ou d'une réplique³³¹.

5.2.3-Récapitulation

D. Parfit affirme que notre existence consiste en l'existence d'un corps, d'un cerveau et d'un ensemble de processus mentaux (mémorisation, poursuite d'intentions, persistance de traits de caractère et de croyances) qui changent graduellement dans le temps. Il explique nos réactions face aux puzzles de l'identité personnelle par une disposition à croire illusoirement que nous sommes une chose simple et analysable qui transcende ces composantes.

Dans les chapitres précédents, nous avons analysé plusieurs facettes de l'expérience consciente qui semblent menacées par son raisonnement : l'intuition d'être une chose présente à la conscience, l'autorité privilégiée sur son propre passé, l'intégration du mental dans un récit de vie et l'unité phénoménale. Tous semblent pointer en direction d'une réalité laissée de côté par l'analyse de D. Parfit, mais nous avons des raisons de penser qu'il y a là une apparence trompeuse.

Dans ce chapitre-ci, nous avons vu une autre facette apparemment irréductible de l'expérience consciente : l'importance. Plusieurs attitudes semblent impossibles à justifier sans nous référer directement à nous-même. D. Parfit voit dans la fiction d'une duplication de soi une démonstration que ce n'est pas le cas. L'identité n'est pas ce qui importe dans la survivance, mais elle accompagne ce qui importe. Ce

³³¹ M. Siderits (2003), p.55-61.

changement de perspective l'incite à dévaluer le rôle du corps dans nos attitudes, à permettre un souci relatif au degré de connectivité psychologique future et à diminuer l'intensité avec laquelle nous le faisons.

La discussion subséquente nous a permis d'élargir notre compréhension des enjeux associés à ce raisonnement. Selon certains, il s'appuie sur la prémisse arbitraire selon laquelle les parties doivent importer plus que le tout et qu'un vaste ensemble de pratiques peuvent être remises en question sur la base d'une exception. Même si ces remarques montrent une limite aux conclusions de D. Parfit, des motivations et précisions additionnelles peuvent lui donner du poids. De nombreuses pratiques ayant un cadre de justifications moindre que la vie, ainsi que l'utilité à prendre les réseaux de pensées et d'expériences fortement intégrées plutôt que le moi comme centre de rationalité vont dans le sens de son raisonnement. Il est donc erroné de décrire la position de D. Parfit comme étant étrangère au contexte de la vie courante.

5.3- Les représentations métaphysiques de la valeur de l'existence

5.3.1- La réalité naturaliste

Le raisonnement de D. Parfit intervient comme un argument parmi d'autres visant à faire pencher des attitudes en conflit dans une direction plutôt que dans l'autre. En résumé, il vise à réorienter les attitudes égocentrées de notre vie en général sur un flux de conscience et autres événements mentaux plus immédiats. De plus, il relativise ainsi l'importance de ce qui nous sépare de la vie des autres (corporalité incluse). Leurs expériences nous semblent inaccessibles, mais elles ne le sont pas tant plus que notre futur lointain. Pourtant, cela ne nous empêche pas de les prendre en considération en y voyant un intérêt. Ces finalités sont en accord avec la vie courante. Nous ne sommes pas régulièrement en train situer notre être en agissant. Nous améliorons spontanément le sort de nos proches, l'horizon d'une vie est rarement le cadre de nos décisions, la direction unique de nos projets résultent d'un

apprentissage et il est utile de se concentrer sur le réseau de pensées et d'expériences qui est fortement connecté à notre situation présente, car c'est sur cela que nous avons du pouvoir.

L'imaginaire d'une fission agit comme une technique de visualisation pour accorder notre théorie à propos de la rationalité avec ces faits tangibles et concrets qui la contredisent. Face à une telle situation, la frontière entre « moi » et « un autre », mais aussi la distance entre « moi » et « le flux de conscience » apparaît plus clairement comme étant un choix de perspective sur une réalité pouvant être décrite sans y faire référence.

Notre existence est externe à notre flux de conscience, alors que ce qui nous importe lui est intuitivement interne. C'est ce que B. Williams et R. Swinburne trouvent inconfortable dans la duplication : elle fait dépendre notre existence d'un « extérieur ». Or, les chapitres précédents visaient à montrer comment nous ne sommes justement pas « internes » à la conscience de nous-mêmes. Cependant, l'intégration de nos pensées et de nos expériences dans un ensemble plus large peut être perçue comme un intérieur dont l'existence ne dépend pas d'un extérieur. C'est pourquoi D. Parfit estime qu'il contient ce qui importe dans la survivance³³².

De plus, il y a une différence d'attitude que nous observons entre le niveau de description impersonnelle et personnelle de nos vies. Méditer sur la trivialité de ce qui les distingue peut faciliter un changement d'attitudes allant en direction du niveau impersonnel, surtout si ce changement est supporté par un ensemble d'autres motivations. Refuser cet argument en brandissant un moi dont la valeur n'a pas besoin d'intermédiaires pour être démontrée fait preuve de dogmatisme.

³³² C'est la réponse que D. Parfit (1984) fait à B. Williams, p.266-273.

Le piège illustré par cette discussion est de présumer qu'étant de nature métaphysique, l'argument de D. Parfit entre en conflit avec un système de justification naturaliste qui repose sur la psychologie humaine et l'immersion dans une culture donnée. Ce n'est pas le cas. L'utilisation de l'imaginaire et l'analyse des composants d'une entité font partie des outils légitimes à employer dans une argumentation et il y a suffisamment de valeurs et de dispositions ambiantes qui peuvent s'y greffer pour aller dans la direction qu'il nous propose sans faire.

5.3.2- La part de vérité et d'illusion essentialiste

La réponse précédente est cependant insatisfaisante. Il y a un enjeu central qu'elle laisse de côté. L'importance de notre existence contient intuitivement un fondement qui transcende notre psychologie et nos normes culturelles observables. Une partie de cette intuition est vraisemblablement illusoire. C'est l'hypothèse qui conclura ce chapitre. Je l'élaborerai en analysant nos réactions à deux fictions : le zombie et la réplique.

5.3.2.1- Zombie et importance de la conscience

Prenons l'existence de la conscience phénoménale. Des auteurs comme T. Nagel et D. Chalmers nient qu'être conscient consiste en un phénomène neurochimique et comportemental observable d'une perspective extérieure³³³. Un des moyens de l'illustrer, c'est d'invoquer une fiction où un être possède l'apparence de la conscience sans l'avoir réellement, c'est-à-dire sans ressentir l'effet que cela lui fait d'être dans le monde et d'y vivre. C'est le zombie. Par exemple, il a un cerveau similaire à nous, puis se comporte de la même manière, mais est aussi vide intérieurement qu'une machine ou un objet inanimé puisse l'être. Nous sommes

³³³ T. Nagel (1981) et D. Chalmers (1996).

intuitivement disposés à imaginer ainsi l'absence de quelque chose. La conscience phénoménale en est la présence³³⁴.

Cette intuition s'accompagne de celle selon laquelle cette conscience est essentiellement valable. Si et seulement si un être la possède, son existence vaut. Imaginons que nos descendants soient séduits par le transhumanisme au point de remplacer leur cerveau biologique par un cerveau électronique plus puissant afin de devenir éternel. Nombreux sont ceux qui n'y voient pas un rêve d'éternité, mais un cauchemar d'extinction de notre espèce. Pourquoi ? Parce que nous y perdriions peut-être la conscience. Nous deviendrions sans nous en apercevoir ces êtres sans capacité de ressentir leur propre existence que nous demande d'imaginer T. Nagel et D. Chalmers. Nous existerions encore de l'extérieur, mais non de l'intérieur³³⁵.

Cette intuition ne vise pas à être cohérente avec les valeurs qui circulent dans notre culture, car en imaginant notre culture muter vers une telle situation, nous ne faisons pas la supposition que de tels êtres changeraient de valeurs. Nous pensons plutôt qu'ils n'en auraient plus. La valeur apparente de leur existence ne correspondrait pas à une valeur réelle désignée par la présence d'une conscience d'être valable. En ce sens, l'importance de la conscience est intuitivement absolue³³⁶.

5.3.2.2- Réplication et importance de l'identité

L'objection adressée par R. Swinburne aux approches réductives selon laquelle elles menacent l'importance de notre existence a la même structure. Il est plus facile de comprendre comment en analysant l'imaginaire d'une réplication de soi.

³³⁴ S. Blackburn (1999), p.67-77 en fait un résumé.

³³⁵ Les réactions décrites à l'endroit de cette fiction sont en conformité avec le fait que, régulièrement dans l'histoire, l'attribution de droit à un être s'accompagnait de la question : a-t-il une âme ?

³³⁶ J'entends par là qu'elle n'est pas relative à une culture ou à un groupe d'individus. Mon raisonnement, c'est qu'une culture ou un groupe d'individus qui n'a pas d'intériorité, mais qui semble seulement en avoir une (exemple : zombies ou androïdes) ne valorise pas réellement des choses et n'a donc pas de valeurs.

Nombreux sont ceux qui imaginent que détruire notre corps et en créer ensuite un autre muni d'états mentaux indiscernables des nôtres est une façon de mourir. D. Parfit reconnaît ce fait et l'illustre par le scénario d'une « ligne embranchée ». Notre réplique est créée quelques heures avant la destruction de notre corps original. Dans la logique de l'indétermination de notre existence, nous pourrions décrire la situation de deux manières. Nous sommes à deux endroits en même temps et deux parties de nous-mêmes communiquent par voie public avant qu'une des deux parties soit détruite sans que le court chevauchement vécu dans le corps original ne soit mémorisé. Nous mourrions en ayant le temps de discuter quelques temps avec un autre être capable de continuer presque parfaitement nos expériences et nos pensées³³⁷. Or, reconnaît D. Parfit, peu de gens y verraient un cas d'indétermination. « I admit that this is one of the cases where my view is hardest to believe³³⁸. » Ils réagiraient comme un mourant et non comme un être muni d'un pouvoir d'ubiquité exposé à une courte perte de mémoire.

Selon lui, cela ne contredit pas sa théorie, mais montre comment nous nous illusionnons sur notre existence. Cette thèse se rapproche de celle du zombie. Une société dont les membres se répliqueraient régulièrement pour devenir éternels perdrait-elle la plupart de ce qui lui donne de la valeur, même si extérieurement elle démontrait les signes du contraire ? Cette fiction serait moins cauchemardesque que celle d'une électrisation complète du cerveau, car de tels êtres continueraient d'être conscients *ponctuellement*, mais ils cesseraient de faire l'expérience d'une vie qui *dure significativement* dans le temps. C'est un cas de zombie *partiel*. C'est cette intuition qu'illustre la ligne embranchée de D. Parfit. L'être original perdrait ce lien de conscience phénoménale durable qui nous importe. Il serait zombifié³³⁹.

³³⁷ D. Parfit (1984), p.287-289.

³³⁸ D. Parfit (1984), p.289.

³³⁹ M. Schechtman (2005), p.6-11 analyse elle aussi l'exemple de la ligne embranchée chez D. Parfit comme étant une manière de contredire l'unité phénoménale des expériences dans le temps. H. Langsam (2001) propose également d'analyser la réflexion de D. Parfit comme une remise en cause

L'intuition, c'est que la valeur de cette phénoménologie est absolue. Une culture qui cesse de la reconnaître perd de la valeur au même titre qu'une espèce zombifiée sans le savoir. Les philosophes traitent rarement notre réaction à cette fiction comme étant relative à notre tempérament ou à notre culture, mais comme étant quelque chose à propos duquel un « sens commun » uniforme doit accorder notre jugement³⁴⁰. Refuser l'indétermination équivaut à porter ce jugement. S'il y a des récalcitrants, c'est parce qu'ils n'ont pas cherché à l'intérieur d'eux-mêmes à la bonne place et non parce qu'ils sont différents du reste de l'humanité. C'est pourquoi ces philosophes argumentent en avançant des exercices d'introspection de plus en plus sophistiqués afin de nous permette tous de mettre le doigt sur cette essence fabuleuse³⁴¹.

Cette conclusion est difficile à démontrer auprès d'eux, car leur immersion dans une culture scientifique rend tabou l'appel au mystère et à la magie. Nous pouvons prédire que, dans ce contexte, la part qu'ils y injectent sera dissimulée, niée, refoulée ou sublimée dans un jargon technique³⁴². Cependant, elle est facile à établir en dehors de ce cercle. Le reste de l'humanité croit massivement en l'existence d'une âme qui survit à la destruction du corps et y voit le signe de leur dignité. C'est l'âme que l'être pieux aurait peur de perdre dans la réplique, c'est-à-dire une essence intérieure source de toute beauté et de toute laideur.

des intuitions du sens commun à propos de la conscience phénoménale. D. Parfit lui-même n'est pas explicite à cet effet, référant plutôt à un sentiment de perdre une « intimité » ou une « profondeur ». J'avance cette interprétation en m'appuyant sur l'hypothèse que son argumentation gagne ainsi en clarté.

³⁴⁰ M. Johnston (1989) est un exemple d'exception.

³⁴¹ Citer un auteur serait laborieux, car la littérature entourant le problème de l'identité du moi à travers le temps dans la philosophie anglo-saxonne contemporaine a largement cette structure.

³⁴² Par exemple, traiter comme le fait T. Nagel (1986) le cerveau comme étant moi muni de deux propriétés irréductibles l'une à l'autre évite de dire qu'il croit en la présence d'une âme dans le cerveau, mais il n'affirme pas quelque chose de bien différent. En fait, il voit l'âme en voyant le cerveau plutôt que de la voir dedans.

5.3.2.3- La non-importance de l'identité revisitée

La réflexion D. Parfit se comprend mieux sous cet angle. Il n'y a rien de tel qu'une expérience consciente intérieure qui dure en moyenne de 70 à 80 ans, ni même quelques jours³⁴³. Lorsque nous mémorisons notre passé, nous le faisons intérieurement exactement de la même manière que nous l'imaginons d'une réplique. Il y a bien une conscience qui dure, mais pas dans le sens d'une conscience qui transcende la simple transmission d'une capacité à se représenter ce passé fidèlement via un support physique³⁴⁴. De ce point de vue, il ne fait effectivement aucune différence en terme de *valeur* que ce support soit un réseau de synapses communiquant par dépolarisation potassique et libération de produits chimiques dans des cuves microscopiques ou par ondes radios entre deux machines munies du pouvoir de détruire et de répliquer des cellules.

Cela ne signifie pas qu'il n'y a pas de conscience phénoménale, mais qu'il n'y en pas une qui *dure*. D. Hume ne niait pas l'existence d'une intériorité, mais en affirmait le caractère extrêmement éphémère. La même chose peut être dite de D. Parfit. En ce sens, l'intuition qui concerne la valeur *absolue* de la conscience phénoménale n'est pas illusoire, mais celle qui concerne la conscience phénoménale *de soi*, l'est³⁴⁵. D. Parfit croit que la conscience phénoménale n'est pas réductible, mais affirme que le moi l'est³⁴⁶. C'est probablement pourquoi il arrête là la réduction de la valeur.

³⁴³ M. Tye (2003) évalue sa durée à une journée, alors que D. Dennett (1991) en nie la présence entre deux clignements d'yeux, p.355-356.

³⁴⁴ Une conscience d'accès.

³⁴⁵ « Soi » étant ici compris comme une unité réfléchie et durable.

³⁴⁶ D. Parfit (1995): « Suppose we are studying some creature which is very unlike ourselves, such as an insect, or some extraterrestrial being. We know all the facts about this creature's behavior, and its neurophysiology. The creature wriggles vigorously, in what seem to be a response to some injury. We ask "Is it conscious, and in great pain? Or is it an insentient machine?" Some behaviorist might say, « That is a merely verbal question. These aren't different possibility, either which might be true. They are merely different description of the same state of affairs ». *That I find incredible*. These description give us, I believe, two quite different possibilities. It could not be an empty or a merely verbal question whether some creature was unconscious or in great pain." D. Parfit (1999), p.258: "I believe

C'est dans ce contexte que nous pouvons interpréter la conclusion de D. Parfit selon laquelle la continuation de l'expérience présente via des souvenirs, des actions, des traits de caractère et des croyances futures ont l'importance que n'a pas l'identité du moi. Reconnaissant la valeur *absolue* de la conscience phénoménale en un instant et le fait qu'elle est bien réelle, il devient important qu'elle s'inscrive dans un ensemble plus large d'autres consciences. Ce qui n'a pas d'importance *absolue*, c'est le *nom* donné à ce réseau et son *nombre* (qu'il y en ait seulement un). Cette importance est cependant moindre que nous le pensons, car elle équivaut à celle d'avoir une famille pour donner de la portée à notre vie. Dans les deux cas, l'importance consiste à donner de l'ampleur à une expérience consciente qui, isolément, serait vide de sens. Cela est valable, mais nous sommes portés à croire que ce n'est pas comparable à survivre. C'est cette addition qui est illusoire selon D. Parfit.

Supposons que l'humanité soit infectée d'une maladie mortelle et que la seule manière d'y répondre est d'implanter dans une machine nos pensées et expériences de manière à créer à tous les jours une nouvelle copie qui en poursuit la saveur. Notre choix est cela ou la disparition pure et simple. Y-a-t-il, intuitivement, une valeur essentielle relative à la *survie* d'effectuer le premier choix ? Intuitivement : oui. Dans ce contexte, nous apparaîtrait-il important que la machine ne reproduise qu'une seule série de répliques afin d'être bien sûr qu'elle puisse être appelée « moi » ? Intuitivement : non³⁴⁷. Ce que D. Parfit affirme, fondamentalement, c'est que la *survie* de la conscience phénoménale en un instant a *toujours* cette forme.

that all experiences have intrinsic qualitative features, and that their qualitative character could be thought about without the thought of the subject for whom they have the character."

³⁴⁷ Bien sûr, je rapporte ma réaction. La réponse de ceux qui me lisent est le moyen de l'infirmier ou de le confirmer.

5.4- Conclusion

D. Parfit solutionne un problème posé à son réductionnisme par la possibilité d'une duplication de soi en affirmant que c'est l'intégration des expériences et des pensées dans des processus comme la mémoire, la poursuite d'intention, la persistance de traits de caractère et de croyances qui importent dans la survivance et non l'identité. La fission illustre comment notre existence dépend d'un fait externe et de l'ajout d'un mot sur une situation préalablement comprise, alors que ce qui nous importe dépend de la réalité et d'un fait interne. Plusieurs auteurs lui reprochent d'accorder arbitrairement l'importance des parties sur le tout, alors que nous pouvons faire l'inverse. Il altère notamment radicalement la perspective adoptée sur nos vies sans justification. De plus, ils estiment préférable de faire de la fission une exception illustrant une règle. Contrairement à ces critiques, des précisions peuvent être apportées concernant les motivations d'une réduction et les éléments plus précis impliqués dans notre existence afin de supporter les conclusions de D. Parfit. J'ai conclu cependant qu'il y a une intuition impossible à préserver et que c'est celle d'importer absolument en vertu d'une conscience phénoménale qui dure.

CONCLUSION

Dans le premier chapitre, j'ai présenté le réductionnisme du moi de D. Parfit comme une théorie de l'erreur niant l'intuition selon laquelle l'observation interne nous révèle une chose qui transcende le corps et les processus mentaux particuliers à laquelle nous correspondrions. Son argumentation fait ressortir aux moyens de fictions imagées (transplants de traces mnésiques, réplication, cogito révisé, spectre de recomposition de soi, examen de physique, fission) comment celle-ci peut se réduire à n'être qu'une projection verbale flexible sur une réalité préalablement comprise. C'est dans ce contexte que le débat entre une conception centrée sur une chose physique singulière et la continuité de processus mentaux est indéterminé, car les deux échouent à rendre compte de cette intuition. Suite à cet échec, nous sommes amenés à réviser notre conception du moi et ces deux approches en sont des candidats possibles, mais aucune des deux ne peut prétendre répondre à l'intuition commune. Le refus de tirer cette conclusion en poursuivant l'appel à des expériences de pensée de plus en plus sophistiquées ou en invoquant des nécessités issues de l'ontologie fondamentale sont des indices selon quoi une conception réductionniste de soi n'est pas encore intégrée de façon cohérente parmi les philosophes et qu'ils ont encore gardé une réserve non réductionniste. Ce fait suggère la présence en nous d'une disposition à l'erreur lorsque nous sommes placés en situation de devoir décrire la conscience que nous avons de notre propre existence.

Dans le second chapitre, j'ai analysé l'impact de l'argumentation de D. Parfit sur l'autorité apparente que nous avons sur notre propre passé. Il cherche à décrire la conscience de nos expériences passées sans faire référence à l'identité du moi à travers le temps afin d'éviter un problème de circularité. Plusieurs lui ont reproché d'enlever ainsi à la mémoire sa spontanéité en exigeant un effort de validation supplémentaire à la vie courante. Nous sommes immunisés contre l'erreur de

d'identification, car notre point de vue implique de ne pas nous chercher nous-mêmes dans la scène que nous nous représentons. Les exemples de D. Parfit pour éviter la circularité ont tous comme trait commun d'altérer cette perspective et donc de détruire ce dont il cherche à faire une analyse. J'ai soutenu que nous pouvons raffiner notre compréhension du contexte rendant possible cette spontanéité sans se référer à soi. Il suffit d'imaginer des conditions raisonnables qui vont disposer nos auditeurs à accepter sans questionnement nos énoncés relativement à la vie des autres sans les induire en erreur. Cela montre comment l'approche de D. Parfit frustre notre intuition relative à la présence d'une autorité privilégiée sur notre propre passé autant que celle de ses critiques, car les deux font reposer notre autorité sur des intermédiaires faillibles. Les seconds spécifient simplement que le contexte nous dispense d'avoir à y penser. À moins de faire intervenir un moi muni d'une propriété mystérieuse, cette attente est illusoire.

Dans le troisième chapitre, j'ai discuté d'objections selon lesquelles D. Parfit ne tient pas compte de la signification holistique, active et durable de l'expérience consciente. Les événements ponctuels signifient quelque chose par le rôle qu'ils jouent dans l'histoire de notre vie et ne peuvent donc être décrits sans y renvoyer. De plus, cette vie est profondément quelque chose que nous faisons et non pas quelque chose qui nous arrive. Nous avons vu que la position de D. Parfit peut s'appuyer sur le fonctionnement du cerveau par systèmes séparés, sur le fait que ces caractéristiques ne sont ni nécessaires, ni suffisantes pour rendre compte des intuitions associées à l'identité du moi et sur la possibilité d'en rendre compte au moyen de pronoms légèrement différents du « je ». J'ai montré comment des faits courants appuient sa démarche. Sur de longues périodes, le contexte original d'un événement devient de plus en plus étranger, nous avons rarement une prise sur notre futur et la phénoménologie de l'agir ne suppose pas notre présence. Ici encore, le problème de ces critiques consiste à négliger la part d'illusion contenu dans leur postulat d'une nécessaire unité de signification à travers le temps. La spontanéité avec laquelle

celle-ci est prise pour acquise en présence de faits lui étant manifestement contradictoires tend à la révéler.

Dans le quatrième chapitre, j'ai réfléchi à l'argumentation de D. Parfit montrant que l'unité des expériences phénoménales peut être décrite sans renvoyer à soi. Il cherche à défaire l'intuition qu'un tel phénomène révèle la présence d'une chose de plus que le corps et que les processus mentaux particuliers. Il s'appuie sur une analyse des cas de cerveaux divisés et une fiction de division brève de la conscience en deux, deux exemples où l'unité des expériences se décrit mieux sans se référer à soi. Des objections montrent que les cas de cerveaux divisés ne manifestent pas des signes de division aussi forte qu'il le présume et qu'ils ne sont pas simultanés. De plus, la description de sa fiction suppose la présence de deux êtres et non d'un seul avec deux consciences, car le moi se définit comme le sujet d'une perspective phénoménale. Or, la division phénoménale demeure une explication plausible des cas empiriques même s'il lui manque la complexité d'une pensée réfléchie. De plus, des cas de stress post-traumatique peuvent recevoir une explication similaire. D'autre part, une définition concurrente de la notion de « moi » référant à une unité durable pouvant occasionnellement renvoyer à elle-même par voie publique rend possible une division de la conscience. En fait, la pulsion forte à nier la présence d'une division lors des cas de cerveaux divisés et la spontanéité avec laquelle des nécessités conceptuelles sont proclamées sans envisager sérieusement les alternatives sont une autre raison de penser que les philosophes qui se posent ces questions ont la croyance que le moi est quelque chose de plus que ce qu'ils admettent dans leur ontologie. Nombreux sont ceux qui croient voir directement leur existence dans l'expérience consciente en pointant un aspect ou un autre, mais ils confondent une théorisation de leur expérience avec ce dont ils font réellement l'expérience. C'est de là que vient plus vraisemblablement l'apparence d'une nécessité.

Dans le cinquième chapitre, j'ai analysé la thèse de D. Parfit selon laquelle l'identité n'est pas ce qui importe dans la survivance. Il solutionne des problèmes posés par la possible duplication de soi aux approches réductionnistes. Celles-ci font dépendre notre existence d'un fait externe et trivial. Plutôt que de conclure en la fausseté de cette approche, il affirme que c'est la présence de l'identité qui n'a pas d'importance, mais qu'elle accompagne seulement ce qui a de l'importance. Ce qui importe est la continuation de notre expérience présente par un ensemble de processus mentaux. Ce critère ne fait pas dépendre ce qui importe d'un fait extérieur. Il raisonne en partant du principe que les éléments d'un tout doivent contenir l'importance du tout. Deux objections ont été analysées. D'une part, le tout peut importer plus que les éléments qui le constituent, surtout lorsqu'il consiste en une perspective spéciale sur le mental. D'autre part, il est plus raisonnable de considérer la fission comme une exception à une règle, car nous ne pouvons pas imaginer ce qui adviendrait de nos pratiques si nous nous rallierions à la conclusion de D. Parfit. J'ai présenté des réponses montrant au contraire que ces conclusions s'accordent avec des motivations courantes et que la perspective qu'elle implique n'est pas étrangère à l'expérience réelle. Cependant, cette discussion néglige qu'il y a bien une dimension contre-intuitive à la conclusion de D. Parfit, mais qu'elle repose plutôt sur la conviction que la conscience est une valeur absolue et que celle-ci s'étend durablement dans le temps. Cependant, cette intuition est illusoire et il y a donc matière à dévaluer une partie de l'importance associée à l'identité du moi à travers le temps.

Toutes ces réflexions ont en commun de porter sur un aspect de soi qui semble désigner une dimension irréductible, impossible à définir sans se viser soi-même comme cible d'un acte d'autoréférence et dont les frontières doivent être fixées de façon déterminée. De nombreuses critiques adressées à D. Parfit cherchent à en rendre compte sans faire intervenir une chose spéciale de plus que le corps, le cerveau et des processus courants comme la mémoire ou la poursuite d'intention. Les réponses que j'ai avancées visent à montrer que cette démarche échoue. En l'absence

d'une telle chose, cette rigidité de la référence à soi dans l'expérience consciente est illusoire. Il est alors plus facile de comprendre comment tel ou tel aspect de la vie mentale peut être décrit adéquatement sans faire référence à soi et à notre identité ou comment l'opposition entre plusieurs conceptions du moi peut être indéterminée. La résistance à tirer cette conclusion suggère la présence d'une croyance non explicite contraire à notre ontologie explicite.

BIBLIOGRAPHIE

Alter, T. "A defense of the necessary unity of phenomenal consciousness" *Pacific Philosophical Quarterly*, 91, 2010: 19-37.

Ayers, M. *Locke*, London, Routledge, 1991.

Bayne T. et Chalmers, D. « What is the unity of consciousness?" Dans A.

Cleeremans (ed.), *The Unity of Consciousness: Binding, Integration and Dissociation*, Oxford, Oxford University Press, 2003: p.23-58.

Bayne T., *The unity of consciousness*, Oxford, Oxford University Press, 2010.

Belzer, M. "Self-Conception and Personal Identity : Revisiting Parfit and Lewis with an eye on the grip of the unity reaction", *Social Philosophy and Policy*, vol. 22, no.2, 2005: p.126-164

Bernecker, S. *Memory: a philosophical study*, New-York, Oxford University Press, 2009.

Bernier, P. « La pensée sans sujet pensant », *Dialogue*, Vol. 49, December 2010 : p. 589-602.

Blackburn, S. "Has Kant refuted Parfit ?", dans *Reading Parfit*, édité par Dancy, J., Oxford, Blackwell Publisher, 1997: p.180-201.

Penser, une irrésistible introduction à la philosophie, Paris, Flammarion, 1999.

Brink, D.O "Rational Egoism and the Separateness of Persons", dans Dancy, J. *Reading Parfit*, Oxford, Blackwell publisher, 1997: p.96-134.

Buford, C. "Memory, Quasi-memory, and Pseudo-quasi-memory", *Australasian Journal of Philosophy*, vol.87, 2009: p.465-478.

Burge, T. "Memory and Persons", *The Philosophical Review*, vol. 112, no.3, 2003: p.289-337.

Cassam, Q. "Parfit on Persons", *Proceedings of the Aristotelian society*, vol. 93, 1993: p.17-37

Self and world, Oxford, Oxford University Press, 1999.

Chalmers, D. J. *The conscious mind: In search of a fundamental theory*, Oxford university press, 1996.

Dainton, B. "The Self and the Phenomenal", *Ratio* vol.17, no. 4 , 2004: p. 365–389

The Phenomenal Selves, Oxford, Oxford University Press, 2008.

Dancy, J. (ed), *Reading Parfit*, Oxford, Blackwell Publisher, 1997.

Dennett, D. *Consciousness explained*, New-York, First Paper Back edition, 1991.

Descartes, R., *Méditations métaphysiques*, Paris, GF Flammarion, 1979.

Evans, G., *The varieties of reference*, Oxford, Clarendon Press, 1982.

Fasching, W. "The mineness of experience", *continental philosophical review*, vol. 42, 2009: p.131-148.

Fisette, D. et Poirier, P., *Philosophie de l'esprit : état des lieux*, Paris, Librairie philosophique Vrin, 2000.

Garrett, B. *Personal Identity and Self-Consciousness*, London, Routledge, 1998.

Gendler, T.S., "Personal Identity and thought-experiments", *The philosophical Quarterly*, vol.52, no.206, 2002: p.34-54.

Ganeri, J. "Cross-Modality and the Self", *Philosophy and Phenomenological Research*, vol. LXI, no.3, 2000: p.639-657.

Gilead, A., "A Humean Argument for Personal Identity", *Metaphysica*, vol. 9, no.1, 2008: p.1-16.

Gilett, G. "Brain Bisection and Personal Identity", *Mind*, vol.94, no.378, 1986: p.224-229.

Hamilton, A. « Mémoire, conscience de soi et identification de soi », *philosophie*, vol.2, no.106, 2010 : p.104-123.

Hofstadter, D. et Dennett, D. (Ed.), *The Mind's I : Fantasies and Reflections on Self and Soul*, New-York, Bantams Books, 1981.

Huemer, M. *The problem of political authority: an examination of the right to coerce and the duty to obey*, New-York, Palgrave Mcmillan, 2013.

Hume, D. *Traité de la nature humaine, Livre I: De l'entendement*, Chicoutimi, Bibliothèque Paul-Émile-Boulet de l'Université du Québec à Chicoutimi, <http://bibliotheque.uqac.ca/> (en ligne)

Howell, R., "Immunity to Error and Subjectivity", *Canadian Journal of Philosophy*, p.37, no.4, 2007: p.581-604

Johnston, M. "Human Concerns without Superlatives Selves", dans J.Dancy *Reading Parfit*, Oxford, Blackwell Publisher, 1997: p.149-179.

Relativism and the Self dans M. Krausz (ed.), *Relativism: Interpretation and Confrontation*. Notre Dame, Notre Dame University Press, 1989.

Klein, S., Nichols, S. "Memory and the Sense of Personal Identity", *Mind*, vol. 121, no. 483, 2012: p.677-702.

Kolak, D. "Room for a view: on the metaphysical subject of personal identity", *Synthese*, 2008, vol.162, no.3: p.341-372.

"The metaphysics and metapsychology of personal identity: why thought experiments matter in deciding who we are." *American Philosophical Quarterly*, 30, no. 1, 1993: p.39-50.

Korsgaard, C.M. "Personal Identity and the Unity of Agency: A Kantian Response to Parfit", *Philosophy & Public Affairs*. Vol. 18, no. 2, 1989: p. 101-132

Langsam, H. "Pain, Personal Identity, and the Deep Further Fact", *Erkenntnis*, vol.54, 2001: p.247-271.

Locke, J. *Enquête sur l'entendement humain*, Paris, Librairie Générale française, 2009.

Traité du gouvernement civil, Paris, GF flammariion, 1984.

Madell, G.C. *The Identity of the Self*, Edinburgh, Edinburgh University Press, 1981.

Martin, R. "Memory, Connecting and what matters in Survival", *Australasian Journal of Philosophy*, vol.65, no.1, 1987: p.82-97.

Self-concern: An experiential approach to what matters in survival, Cambridge, Cambridge University Press, 1998.

McDowell, J. "Reductionism and the First Person", *Reading Parfit* ed. Par Dancy, J., Oxford, Blackwell publisher, 1997.

Nagel, T. "Brain Bisection and the Unity of Consciousness", *Synthese*, vol. 22, no.3/4, 1971: p.396-413

The View from nowhere, Oxford, Oxford University Press, 1986

"What it is like to be a bat" dans *The minds's I: fantasies and reflection on Self and Soul* édité par D. R. Hofstadter et D. Dennett, New-York, Bantam book, 1981.

Ninan, D. "Persistence and the First-Person Perspective", *Philosophical Review*, 2009: p.425-464.

Noonan, H. *Personal Identity*, London, Routledge, 1989

Northoff, G. "Are "q-memories" empirically realistic?: A neurophilosophical approach", *Philosophical Psychology*, vol.13, no.2, 2000: p.191-211

Olson, E., *What are we*, Oxford, Oxford University Press, 2007

Parfit, D. *Reasons and Persons*, Oxford, Clarendon Press, 1984.

"The Unimportance of Identity", 1995, http://www.stafforini.com/txt/parfit_-_unimportance_of_identity.pdf (en ligne)

"Experiences, Subjects and Conceptual Schemes", *Philosophical Topics*, vol.26, no. 1 et 2, 1999: p.217-270.

Perry, J. *Personal Identity*, Berkeley, Los Angeles, London, University of California Press, 1975

Proust, J. "Thinking of oneself as the same", *consciousness and cognition*, vol. 12, no. 4, 2003: 495-509.

Pryor, J. "Immunity to Error through Misidentification", *philosophical topic*, vol. 26, no.1, 1999: p.271-304.

Reid, M.D. "Memory as Initial Experiencing of the Past", *Philosophical Psychology*, vol. 18, no.6, 2005: p.671-698.

Roache, R. "A defense of quasi-memory", *Philosophy*, vol. 81, no.2, 2006: p.323-355.

Rothbard, M. *For a new liberty: the libertarian manifesto*, Auburn, L.V. Mises institute, 2006.

Rovane, C. "Branching Self-Consciousness", *The Philosophical Review*, vol. 99, no.3, 1990: p. 355-395.

The Bounds of Agency: An Essay in Revisionary Metaphysics, Princeton University Press, 1998.

"Self-reference: the radicalisation of Locke", *The journal of philosophy*, vo.90, no.2, 1993: p.73-97.

Rudd, A., "Narrative, Expression and Mental substance", *Inquiry*, vol. 48, no.5, 2005: p.413-435.

Sartre, J-P., *La Transcendance de l'ego*, Paris, Vrin, 1992.

Schechtman, M., "Personhood and Personal Identity", *The Journal of Philosophy*, vol. 87, no.2, 1990: p. 71-92

"Experience, agency and personal identity", *Social Philosophy and Policy*, vol. 22, no.2, 2005, p. 1-24

Shoemaker, D. "Theoretical Persons and Practical Agents", *Philosophy & Public Affairs*, vol.25, no.4, 1996: p.318-332.

"Personal identity and ethics", *Stanford encyclopedia of philosophy*, 2008, <http://www.science.uva.nl/~seop/archives/win2008/entries/identity-ethics/> (en ligne)

"Selves and Moral Units », *Pacific Philosophical Quarterly*, vol. 80. No 4, 1999: p.391-419.

Shoemaker, S. "Persons and Their Pasts", *Philosophical Quarterly*, vol.7, no. 4. 1970: p.269-285.

Shoemaker S. et Swinburne, R., *Personal Identity*, Oxford, B.Blackwell, 1984.

Siderits, M. *Personal identity and Buddhist philosophy: empty persons*, Hampshire, Ashgate Publishing limited, 2003.

Slors, M. "Personal Identity, Memory, and Circularity: An Alternative for Q-Memory", *The Journal of Philosophy*, vol.98, no.4, 2001: p. 186-214.

Strawson, G. "The self and the SESMET", *Journal of Consciousness Studies*, vol.6, no. 4, 1999: p.99-135

Stokes, P., "Uniting the perspectival subject: Two approaches", *Phenomenologie and Cognitive Sciences*, vol.10, 2010: p. 23–44.

Swinburne, R. G. "Personal Identity", *Proceedings of the Aristotelian Society*, Vol. 74, (1973 - 1974): p. 231-247.

Tye, M. *Consciousness and persons: Unity and identity. Representation and mind*. Cambridge, MIT Press, 2003.

Unger, P. *Identity, Consciousness and Value*, Oxford, Oxford University Press, 1990.

Velleman, J.D. "Self to Self", *The Philosophical Review*, vol.105, no. 1, 1996: p.39-76.

Williams, B. "The Self and the Future", *The Philosophical Review*, vol.79, no.2, 1970, p.161-180.

The problem of the Self, Cambridge University Press, 1973.

Wolf, S. "Self-Interest and Interest in Selves", *Ethics*, vol.96, no.4, 1986, p.704-720.

Wollheim, R., *The thread of life*, Cambridge, Harvard University Press, 1984.

Zahavi, D., "Self and Consciousness", *Exploring the Self*, ed. Par Zahavi, D., Amsterdam-Philadelphia, John Benjamins Publishing Company, 2000: p.55-74.

Zahavi, D. et Parnas, J., "Phenomenal consciousness and self-awareness: a phenomenological critique of representational theory", *Journal of Consciousness Studies*, vol.5, no.5–6, 1998: p.687–705.